

# Exponential estimates of slow manifolds

K. Uldall Kristiansen and C. Wulff

Department of Mathematics,  
Technical University of Denmark,  
2800 Kgs. Lyngby,  
DK  
and

Department of Mathematics,  
University of Surrey,  
Guildford, GU2 7XH  
UK

## Abstract

In this paper we prove exponential estimates of slow manifolds in analytic systems. The results are obtained for general slow-fast systems and finite-dimensional Hamiltonian systems. For general systems we consider the motion in a Banach space with an unbounded fast vector-field, while for Hamiltonian systems we consider finitely many fast and slow variables. We will prove some conjectures of MacKay from a 2004 reference, and the methods we use are based upon the ideas presented in this paper. The method does not notice resonances, and therefore we do not pose any restrictions on the motion normal to the slow manifold other than it being fast and analytic.

# 1 Introduction

Singularly perturbed systems involving different time and/or space scales arise in a wide variety of scientific problems. Important examples include: meteorology and short-term weather forecasting [15, 16, 26], molecular physics and the Born-Oppenheimer approximation [20], chemical enzyme kinetics and the Michaelis-Menten mechanism [21], predator-prey and reaction-diffusion models [22], the evolution and stability of the solar system [13, 14] and in the modeling of tethered satellites [27, 28]. These systems can also be “artificially constructed” by a partial scaling of variables near a bifurcation [25]. The main advantage of identifying slow and fast variables is dimension reduction by which all the fast variables are “slaved” to the slow ones through the *slow manifold*. Dimension reduction is one of the main aims and tools for a dynamicist and the elimination of fast variables is very useful in for example numerical computations. In fact, since fast variables require more computational effort and evaluations, this reduction often bridges the gap between tractable and intractable computations. An example of this is the long time (*Gyears*) integration of the solar system, see [13, 14]. See also [3, 7, 31] for a numerical treatment of slow-fast systems.

## A conjecture by MacKay.

In [17] R. S. MacKay conjectured on exponential estimates of slow manifolds in analytic systems. It was sketched how slow manifolds with non-degenerate normal behavior could be improved and it was then claimed that an iteration would lead to slow manifolds where the vector-field would have an exponentially small angle to the tangent-space. To explain the method and introduce slow-fast systems let us consider the system

$$\partial_t w = \epsilon W(w, z), \quad \partial_t z = Z(w, z), \quad (1.1)$$

with  $\epsilon$  a small parameter. The analytic vector-fields  $W$  and  $Z$  will in general also depend upon  $\epsilon$ , but we shall suppress this dependency throughout. We assume that  $Z(w, z) = 0$  has a solution  $z = \zeta(w)$ , and introduce  $z_0$  through  $z = \zeta + z_0$  to transform these equations into

$$\partial_t w = \epsilon W_0(w, z_0), \quad \partial_t z_0 = Z_0(w, z_0) = \rho_0(w) + A_0(w)z_0 + R_0(w, z_0), \quad (1.2)$$

with

$$\begin{aligned} \rho_0 &= -\epsilon \partial_w \zeta W(w, \zeta), \\ A_0(w) &= \partial_z Z(w, \zeta) - \epsilon \partial_w \zeta \partial_z W(w, \zeta), \\ R_0(w, z_0) &= \int_0^1 (1-s) \left( \partial_z^2 Z(w, \zeta + sz_0) - \epsilon \partial_w \zeta \partial_z^2 W(w, \zeta + sz_0) \right) z_0^2 ds, \end{aligned}$$

and  $W_0(w, z_0) = W(w, \zeta + z_0)$ . Here  $\partial_w$  is used to denote the (Frechet) partial derivatives  $\frac{\partial}{\partial w}$ , and we will continue to use this symbol regardless of what object is being differentiated. In slow-fast systems one often connects (1.2) with the system

$$\partial_\tau w = W_0(w, z_0), \quad \epsilon \partial_\tau z_0 = \rho_0(w) + A_0(w)z_0 + R_0(w, z_0), \quad (1.3)$$

related to (1.2) by the scaling

$$\tau = \epsilon t. \quad (1.4)$$

If we formally set  $\epsilon = 0$  in (1.2) and (1.3), then two limit systems are obtained. In (1.3), the formal limit is singular leading to the algebraic equation:  $A_0(w)z_0 + R_0(w, z_0) = 0$  (since  $\rho_0 = O(\epsilon)$ ). Due to the singular nature of this limit one also often refers to the theory as singular perturbation theory. The set of points  $M_0 = \{z_0 = 0\}$  satisfying these equations is called a *slow manifold* (or critical manifold) and the corresponding system:  $\partial_\tau x = W_0(w, 0)$  is called *the slow subsystem*. On the other hand, in (1.2), the formal limit leads to *the fast sub-*, or *frozen*, system:  $\partial_t z_0 = A_0(w)z_0 + R_0(w, z_0)$  with  $x$  now considered as a parameter. In this system,  $M_0$  is a set of equilibria and the linearization  $\partial_\tau \delta z_0 = A_0(w)\delta z_0$  about these determine the classification of the slow manifold. In particular, if  $z_0 = 0$  is an elliptic or hyperbolic equilibrium, then the slow manifold  $M_0$  is said to be normally elliptic respectively hyperbolic at the point  $(w, 0)$ .

One of the main tasks in singular perturbation theory is to determine the fate of the slow manifold  $M_0$  for  $\epsilon > 0$  but small, and thus connect the apparent two different limit systems for  $\epsilon \neq 0$ . When  $M_0$  is hyperbolic then there exists a perturbed, invariant slow manifold  $M(\epsilon)$  for  $\epsilon \neq 0$  nearby [4, 5]. For general non-Hamiltonian systems hyperbolicity is generic. On the other hand, normally elliptic slow manifolds, which are generic in Hamiltonian systems, are unlikely to persist because typical perturbations are believed to destroy them [17, 18]. For these types of slow manifolds, one therefore usually aims for something less: *almost invariance*. For singular perturbed Hamiltonian systems with only one fast degree of freedom and an analytic Hamiltonian function, Gelfreich and Lerman [8] showed, using an averaging procedure, the existence of an almost invariant slow manifold nearby. By “almost” it is understood that the error field (the normal component of the vector field restricted to the slow manifold, [17]) is of order  $\mathcal{O}(e^{-C/\epsilon})$ ,  $C > 0$ . For the Hamiltonian example

$$H = \frac{1}{2}x^2 + \frac{1}{2}y^2 + v + \epsilon y f(u), \quad (1.5)$$

with  $f(u) = \sum_{n=1}^{\infty} e^{-n} \sin(nu)$  and  $\omega = dx \wedge dy + \epsilon^{-1} du \wedge dv$ , Neishtadt showed that the slow manifold cannot be improved beyond such an estimate, see [8]. The exponential estimate is therefore the best one can aim for in a general setting for normally elliptic slow manifolds.

The method of averaging used in [8] does not extend to several fast variables primarily due to the general lack of control of resonances between the fast variables.

The averaging method of [8] aims at more than we do: the results of [8] do not only provide exponential estimates of a slow manifold, they also provide a  $\mathcal{O}(1)$ -foliation, parametrized by the action variable, of almost invariant slow manifolds. The method therefore, in some sense, also addresses stability, not only existence of the slow manifold. The reference [19] extends the results of [8] to infinite dimensional slow dynamics. The results of [19] hold true for *spatially* Gevrey smooth solutions, which allow for a Galerkin approximation that separates the vector-field into a bounded one and an exponential small remainder. Matching up the error from the averaging procedure with the error from the Galerkin approximation the authors obtained a slow manifold with  $\mathcal{O}\left(\exp\left(-C\epsilon^{-\frac{p}{p+1}}\right)\right)$  error field, where  $p$  is a positive parameter depending on the Gevrey space. Their results hold true for both Hamiltonian

and general systems. The references [29, 30] also provide exponential estimates of particular slow manifolds in geophysical models by obtaining optimal truncations of the “super-balance equation” (invariance equation) of Lorenz [16]. An extension to more general systems is, however, not provided.

### MacKay’s method for improving a slow manifold.

The method of MacKay does not separate normally hyperbolic from normally elliptic slow manifolds. The general assumption is just that  $\|A_0(w)^{-1}\|$  is bounded so that the normal motion is truly fast. Setting  $\partial_t z_0 = 0$  in (1.2) therefore gives, by applying the implicit function theorem (bearing in mind that  $R_0$  is quadratic in  $z_0$ ), a solution  $z_0 = \zeta_0(w)$  close to  $A_0(w)^{-1}\rho_0(w)$ . The graph  $z_0 = \zeta_0(w)$  will be the improved slow manifold. To show that this is indeed an improved slow manifold, one then straightens out the new slow manifold by introducing  $z_1$  through  $z_0 = z_1 + \zeta_0$ . Then the equations become

$$\partial_t w = \epsilon W_1(w, z_1), \quad \partial_t z_1 = Z_1(w, z_1) = \rho_1(w) + A_1(w)z_1 + R_1(w, z_1),$$

with

$$\rho_1 = -\epsilon \partial_w \zeta(w) W_0(w, \zeta_0), \tag{1.6}$$

and so, as  $\zeta_0 = \mathcal{O}(\epsilon)$ , we have  $\rho_1 = \mathcal{O}(\epsilon^2)$  which is the measure of the error field, an improvement from  $\mathcal{O}(\epsilon)$  to  $\mathcal{O}(\epsilon^2)$ , so  $M_1 = \{z_1 = 0\}$  is an improved slow manifold. MacKay’s method, though viewed slightly differently, is actually identical to the method suggested by Fraser [6]. An asymptotic analysis of this method is also given in [12], they do, however, only show formal estimates and do not obtain exponential estimates. Furthermore, they do not consider the Hamiltonian case. We note that  $\rho_1$  actually vanishes at a true equilibrium where  $W_0(w, \zeta_0(w)) = 0$ , and the improved slow manifold  $M_1 = \{z_1 = 0\}$  therefore includes all nearby equilibria. One can continuously perform this procedure as the order of differentiability of  $X$  and  $Y$  allows, obtaining a slow manifold with an  $\mathcal{O}(\epsilon^n)$  error field, but this estimate is not in general uniform in  $n$ . For analytic systems, however, MacKay conjectured that one could obtain an  $\mathcal{O}(e^{-C/\epsilon})$  estimate by applying Neishtadt-type estimates [23, 24], which have also successfully been applied in [19, 8]. This is the first main result, which we present formally in Theorem 3.1 and prove in section 4. In fact, MacKay believed one could obtain a stronger result. He conjectured, we guess based on the expression for  $\rho_1$  in (1.6), that the error could be estimated point wise in  $x$  at each step so that formally an estimate of the form  $\mathcal{O}((\epsilon \|W_n(w, 0)\|^n)$  could be obtained leading to an estimate  $\mathcal{O}(e^{-C/(\epsilon \|W(w, \zeta_*(w))\|)})$  with  $n = \mathcal{O}(\epsilon^{-1})$  where  $z = \zeta_*(w)$  is the improved slow manifold. We will show that this is wrong by considering the following counter example:

**Example 1.1** We consider the simple two-dimensional example:

$$\dot{w} = \epsilon W(w, z) = \epsilon f(w), \quad \dot{z} = Z(w, z) = \epsilon w - z. \tag{1.7}$$

Here  $z = 0$  is actually normally hyperbolic and there is a true invariant manifold nearby but this is irrelevant for the purpose of illustrating why MacKay’s conjecture is in general incorrect. The system has an equilibrium, if  $f(w_e) = 0$ , given by  $(w_e, \epsilon w_e)$ . We shall assume

that the root of  $f$  is simple so that  $f'(w_e) \neq 0$ . Note that (1.7) is already of the form (1.3), i.e.,  $z = z_0$ ,  $Z = Z_0$ . The first step of the procedure is then performed by solving  $Z_0(w, z_0) = 0$  with respect to  $z_0$  giving  $z_0 = \zeta_0(w) = \epsilon w$ . Therefore  $z_1 = z - \zeta_0$  solves

$$\dot{z}_1 = Z_1(w, z_1) = \rho_1 - z_1,$$

with

$$\rho_1 = Z_1(w, 0) = -\epsilon^2 f(w).$$

But at the next step when solving  $Z_1(w, z_1) = 0$  for  $z_1 = \zeta_1(w)$  we obtain

$$\zeta_1(w) = \rho_1 = -\epsilon^2 f(w),$$

and so

$$\dot{z}_2 = \dot{z}_1 - \dot{\zeta}_1 = \rho_1 - (z_2 + \rho_1) + \epsilon^3 f'(w) f(w) = \epsilon^3 f'(w) f(w) - z_2.$$

Notice that  $\rho_2 = \epsilon^3 f'(w) f(w)$  cannot be bounded from above by an expression with  $|f(w)|^2$  as a factor. This is due to the fact that  $w_e$  is only a single root of the dominant order  $\epsilon^3$  term of  $\rho_2$ . Applying the procedure  $n$  times will in general introduce  $f^{(n-1)}$  in the expression for  $\rho_n$ . The corresponding term will have a factor of  $f$ . This is basically just a result of the Leibnitz rule from the continuing differentiation of the  $\zeta_n$ 's.  $\square$

We remark that our results on exponential estimates (Theorems 3.1 and 3.2 below) also hold true for normally hyperbolic slow manifolds. Since Fenichel's theorems actually guarantee the existence of a slow manifold in this case these estimates may seem somewhat unsatisfactory. However, convergence would imply that the slow manifold would be analytic, which is not always the case [2].

## Hamiltonian systems.

Normally elliptic slow manifolds are believed to be particularly interesting in Hamiltonian systems as stability here is associated with oscillatory normal behavior. In Hamiltonian systems there are invariant manifolds that are not normally hyperbolic. The most obvious examples are regular energy levels and KAM-tori in near-integrable Hamiltonian systems. These invariant manifolds are however non-symplectic. Here we are interested in symplectic slow manifolds on which we can define a "slow" Hamiltonian system.

MacKay, still in [17], suggested a separate method for improving slow manifolds in Hamiltonian systems. The proposed method was described as follows: Consider a Hamiltonian  $H = H(p)$  with symplectic form  $\omega$  and a slow manifold  $M_0$ . Do the following:

- Compute an orthogonal symplectic foliation  $F_p$  so that for every  $p \in M_0$  and

$$\omega(z, w) = 0, \quad z \in T_p F_p, \quad w \in T_p M_0.$$

- Let  $H_p = H|_{F_p}$  and solve this for a nearby critical point  $p_1 = p_1(p)$ .

- Put  $M_1 = \{p_1(p)\}$  and  $\omega_{M_1} = \omega|_{M_1}$ .

Then  $(H|_{M_1}, \omega_{M_1})$  is an improved slow system. However, we believe that this method has some drawbacks. First of all, the method requires the computation of a new slow symplectic form at each step. In fact, we believe that the reason for suggesting an alternative to the general approach in the first place, is that one wishes to introduce transformations that preserve the symplectic structure. Moreover, MacKay's method also requires the computation of orthogonal symplectic foliations at each step.

We will therefore suggest an alternative method that circumvent these issues. Our method is basically a symplectic extension of the general approach outlined above. We will straighten out the improved manifold at each step, ensuring that the transformation involved in this procedure is symplectic. Here we will make use of a Lemma 2 in [8]. The slow symplectic form  $\omega|_M$  will therefore remain constant throughout the iteration. The result we obtain is a symplectic slow manifold with exponentially small error field containing an initially nearby equilibrium, as was also conjectured by MacKay. This is the second main result of the paper. We will present this formally in Theorem 3.2 and prove it in section 5. In section 6 we present a dynamical consequence of our result on the persiststence of a slow manifold with exponentially small gaps.

## 2 Notation, assumptions and preliminaries

Let  $(\mathcal{W}, \|\cdot\|_{\mathcal{W}})$  and  $(\mathcal{Z}, \|\cdot\|_{\mathcal{Z}})$  be real Banach spaces and  $\mathcal{W}_{\mathbb{C}} = \mathcal{W} \oplus i\mathcal{W}$  resp.  $\mathcal{Z}_{\mathbb{C}} = \mathcal{Z} \oplus i\mathcal{Z}$  their complexifications with norms  $\|w_1 + iw_2\|_{\mathcal{W}_{\mathbb{C}}} = \|w_1\|_{\mathcal{W}} + \|w_2\|_{\mathcal{W}}$  and  $\|z_1 + iz_2\|_{\mathcal{Z}_{\mathbb{C}}} = \|z_1\|_{\mathcal{Z}} + \|z_2\|_{\mathcal{Z}}$ . We will from now on denote all norms, including operator norms, by  $\|\cdot\|$ . Hopefully it will be clear from the context what norm is used. Then  $f : \mathcal{V}_{\mathbb{C}} \rightarrow \mathcal{Z}_{\mathbb{C}}$ , with  $\mathcal{V}_{\mathbb{C}}$  an open subset of  $\mathcal{W}_{\mathbb{C}}$ , is analytic if it is continuously differentiable, i.e., if there exists a continuous derivative  $\partial_w f : \mathcal{V}_{\mathbb{C}} \rightarrow \mathcal{L}(\mathcal{W}_{\mathbb{C}}, \mathcal{Z}_{\mathbb{C}})$ , where  $\mathcal{L}(\mathcal{W}_{\mathbb{C}}, \mathcal{Z}_{\mathbb{C}})$  is the Banach space of complex linear operators from  $\mathcal{W}_{\mathbb{C}}$  to  $\mathcal{Z}_{\mathbb{C}}$  equipped with the operator norm, satisfying the following condition

$$\|f(w+h) - f(w) - \partial_w f(w)(h)\| = \mathcal{O}(\|h\|^2).$$

By real analytic we will mean analytic and real when the arguments are real. The higher order derivatives can be defined inductively and  $\partial_w^n f$  becomes a map

$$\partial_w^n f : \mathcal{V}_{\mathbb{C}} \rightarrow \mathcal{L}^n(\mathcal{W}_{\mathbb{C}}, \mathcal{Z}_{\mathbb{C}}),$$

from  $\mathcal{V}_{\mathbb{C}}$  into the Banach space  $\mathcal{L}^n(\mathcal{W}_{\mathbb{C}}, \mathcal{Z}_{\mathbb{C}})$  of all bounded,  $n$ -linear maps from  $\mathcal{W}_{\mathbb{C}} \times \cdots \times \mathcal{W}_{\mathbb{C}}$  ( $n$  times) into  $\mathcal{Z}_{\mathbb{C}}$ . See [11, Appendix A] for a reference on analytic function theory in Banach spaces. When  $\mathcal{V}$  is an open subset of  $\mathcal{W}$  and  $\chi > 0$  then, as in [8], we define  $\mathcal{V} + i\nu$  to be the open complex  $\nu$ -neighborhood of  $\mathcal{V}$ :

$$\mathcal{V} + i\nu = \{w \in \mathcal{W}_{\mathbb{C}} \mid d_{\mathcal{W}_{\mathbb{C}}}(w, \mathcal{V}) < \nu\},$$

where  $d_{\mathcal{W}_{\mathbb{C}}}$  is the metric induced from the Banach norm  $\|\cdot\|$ . In the following let  $\mathcal{B}_r^{\mathcal{Z}}(z) \subset \mathcal{Z}_{\mathbb{C}} = \{u \in \mathcal{Z}_{\mathbb{C}}, \|u - z\| < r\}$  denote a  $\mathcal{Z}_{\mathbb{C}}$ -open ball of radius  $r > 0$  around  $z$  in the Banach space  $\mathcal{Z}_{\mathbb{C}}$ . We frequently need the following Cauchy estimate:

**Lemma 2.1** Assume that  $f : \mathcal{V}_{\mathbb{C}} \rightarrow \mathcal{Z}_{\mathbb{C}}$  is analytic and that  $f$  is bounded on  $\mathcal{B}_{\nu}(w_0) \subset \mathcal{V}_{\mathbb{C}}$  for some  $\nu > 0$ . Then

$$\|\partial_w f(w_0)\| \leq \frac{\sup_{w \in \mathcal{B}_{\nu}(w_0)} \|f(w)\|}{\nu}. \quad (2.1)$$

□

**Remark 2.1** Let  $f : \mathcal{V} + i\nu \rightarrow \mathcal{Z}_{\mathbb{C}}$  be analytic and bounded and  $\nu > \nu_0 > 0$ . Then we can apply this estimate to any  $w_0 \in \mathcal{V} + i(\nu - \nu_0)$  to obtain:

$$\sup_{w_0 \in \mathcal{V} + i(\nu - \nu_0)} \|\partial_w f(w_0)\| \leq \frac{\sup_{w \in \mathcal{V} + i\nu} \|f(w)\|}{\nu_0},$$

which we will write compactly as

$$\|\partial_w f\|_{\nu - \nu_0} \leq \frac{\|f\|_{\nu}}{\nu_0}.$$

This is the form of Cauchy's estimate that we will be using. Similarly, we will by  $\|\cdot\|_{\nu, \sigma}$  denote the sup-norm taking over the domain  $(\mathcal{V} + i\nu) \times (\mathcal{S} + i\sigma)$ . □

We will also use the following generalized version of Taylor's theorem [11]:

**Lemma 2.2** If  $f : \mathcal{V}_{\mathbb{C}} \rightarrow \mathcal{Z}_{\mathbb{C}}$  is  $n$  times continuously differentiable,  $n \geq 1$ , and if the segment  $w + sh$ ,  $0 \leq s \leq 1$ , is contained in  $\mathcal{V}_{\mathbb{C}}$ , then

$$\begin{aligned} f(w + h) &= f(w) + \partial_w f(w)(h) + \cdots + \frac{1}{(n-1)!} \partial_w^{n-1} f(w) h^{n-1} \\ &\quad + \int_0^1 \frac{(1-s)^{n-1}}{(n-1)!} \partial_w^n f(w + sh)(h, \dots, h) ds. \end{aligned} \quad (2.2)$$

The integral remainder is bounded by  $\frac{\|h\|^n}{n!} \sup_{0 \leq s \leq 1} \|\partial_w^n f(w + sh)\|$ . □

Here, we write for  $m \in \mathbb{N}$ ,

$$\partial_w^m f(w) h^m = \partial_w^m f(w)(h, \dots, h).$$

### Assumptions on general system.

By straightening out the slow manifold  $z = \zeta(w)$ , the general slow-fast system (1.1) was transformed into (1.2), repeated here for convenience,

$$\partial_t w = \epsilon W_0(w, z_0), \quad \partial_t z_0 = \rho_0(w) + A_0(w)z_0 + R_0(w, z_0),$$

and we will now state our assumptions on this system:

(G1) Let  $\mathcal{V} \subseteq \mathcal{W}$  be open and let  $\nu_0 > 0$ . We assume that  $A_0(\cdot)^{-1} : (\mathcal{V} + i\nu_0) \rightarrow \mathcal{L}(\mathcal{Z}_{\mathbb{C}}, \mathcal{Z}_{\mathbb{C}})$  is real analytic with  $\|A_0(w)^{-1}\|_{\nu_0} \leq \frac{K_0}{2}$ ;



(G2) Let  $0 \in \mathcal{S} \subseteq \mathcal{Z}$  be open and bounded, with  $C_z > 0$  such that  $\|z_0\| \leq C_z$  for all  $z_0 \in \mathcal{S} + i\sigma_0$  and let  $\sigma_0 > 0$ . Assume that  $\rho_0 : (\mathcal{V} + i\nu_0) \rightarrow \mathcal{Z}_{\mathbb{C}}$   $W_0 : (\mathcal{V} + i\nu_0) \times (\mathcal{S} + i\sigma_0) \rightarrow \mathcal{W}_{\mathbb{C}}$  are real analytic. Furthermore, we assume that the norm of  $W_0$  is uniformly bounded by  $C_{W_0}$  on  $(\mathcal{V} + i\nu_0) \times (\mathcal{S} + i\sigma_0)$ , and

$$\delta_0 = \|\rho_0\|_{\nu_0} = \mathcal{O}(\epsilon), \quad (2.3)$$

where  $\epsilon > 0$  is small;

(G3)  $R_0(\cdot, \cdot) : (\mathcal{V} + i\nu_0) \times (\mathcal{S} + i\sigma_0) \rightarrow \mathcal{Z}_{\mathbb{C}}$  is real analytic and quadratic in  $z_0$  so that  $R_0(w, z_0) = \sum_i \sum_{j \geq 2} a_{ij} w^i z_0^j$ , with proper interpretation of  $a_{ij}$  as multi-linear operators. Furthermore the norm of  $R_0$  is assumed to be uniformly bounded by  $C_{R_0}$  on  $(\mathcal{V} + i\nu_0) \times (\mathcal{S} + i\sigma_0)$ ;

We now write  $\epsilon W_0$  as

$$\epsilon W_0 = (C_{W_0} \epsilon)(C_{W_0}^{-1} W_0), \quad (2.4)$$

and let  $C_{W_0} \epsilon$  be our new small parameter, which we continue to denote by  $\epsilon$ , and replace  $W_0$  by  $C_{W_0}^{-1} W_0$ . We will also continue to denote the new  $W_0$  by the same symbol, hoping that this will not cause unnecessary confusion. We therefore continue with

$$C_{W_0} = 1. \quad (2.5)$$

### Assumptions on Hamiltonian system.

In section 3.2 we will consider a real analytic Hamiltonian system  $H = H(w, z)$ , where  $w = (u, v) \in \mathcal{W}$ ,  $\dim \mathcal{W} = 2d_{\mathcal{W}}$ , are the slow variables and  $z = (x, y) \in \mathcal{Z}$ ,  $\dim \mathcal{Z} = 2d_{\mathcal{Z}}$ , are the fast variables, with  $d = d_{\mathcal{W}} + d_{\mathcal{Z}}$  degrees of freedom,  $d_{\mathcal{W}}$  slow- and  $d_{\mathcal{Z}}$  fast ones, and the symplectic form

$$\omega = dx \wedge dy + \epsilon^{-1} du \wedge dv \quad (2.6)$$

(see also Neisthardt's Example (1.5)). As for the general case above we consider a slow manifold  $M_0$  as a set of constrained equilibria:

$$M_0 = \{z = \zeta(w) = (\zeta^x(u, v), \zeta^y(u, v)) | \partial_z H(w, z) = 0\}.$$

This will be our (first) slow manifold which we aim to improve upon. We will straighten  $z = \zeta(w)$  out by introducing new coordinates through the generating function:

$$G_0(u, v_0, x, y_0) = \langle x, y_0 \rangle + \epsilon^{-1} \langle u, v_0 \rangle + g_0(u, v_0, x, y_0),$$

with

$$g_0(u, v_0, x, y_0) = -\langle \zeta^x(u, v_0), y_0 \rangle + \langle \zeta^y(u, v_0), x \rangle. \quad (2.7)$$

Here  $\langle a, b \rangle = a_1 b_1 + \dots + a_n b_n$  for every pair  $a = (a_1, \dots, a_n), b = (b_1, \dots, b_n) \in \mathbb{C}^n$ . Then  $(u, v, x, y) \mapsto (u_0, v_0, x_0, y_0)$ , for  $\epsilon$  sufficiently small, is a symplectic transformation given implicitly by the equations:

$$\begin{aligned} x_0 &= \partial_{y_0} G_0 = x - \zeta^x, & y &= \partial_x G = y_0 + \zeta^y, \\ u_0 &= \epsilon \partial_{v_0} G_0 = u + \epsilon \partial_{v_0} g, & v &= \epsilon \partial_u G_0 = v_0 + \epsilon \partial_u g. \end{aligned}$$



This transforms  $H$  into

$$H_0(w_0, z_0) = H(w, z) = h_0(w_0) + \langle \rho_0(w_0), z_0 \rangle + \frac{1}{2} \langle A_0(w) z_0, z_0 \rangle + r_0(w_0, z_0), \quad (2.8)$$

with  $\rho_0 = \mathcal{O}(\epsilon)$  on  $(\mathcal{V} + i\nu_0) \times (\mathcal{S} + i\sigma_0)$  where  $\mathcal{V}, \mathcal{S}$  are as before and  $\nu_0, \sigma_0 > 0$ . We now assume the following:

- (H1)  $A_0 = \partial_{z_0}^2 H_0|_{z_0=0} : \mathcal{S} + i\sigma_0 \rightarrow \mathcal{L}(\mathcal{Z}, \mathcal{Z})$  is a symmetric matrix valued real analytic function satisfying  $\|A_0\|_{\nu_0} \leq C_{A_0}$  and  $\|A_0(w)^{-1}\|_{\nu_0} \leq \frac{K_0}{2}$ ;
- (H2)  $\rho_0 = \partial_{z_0} H_0|_{z_0=0} : \mathcal{V} + i\nu_0 \rightarrow \mathcal{L}(\mathcal{Z}, \mathbb{C})$  and  $h_0 = H_0|_{z_0=0} : \mathcal{V} + i\nu_0 \rightarrow \mathbb{C}$  are real analytic functions and  $\delta_0 = \|\rho_0\|_{\nu_0} = \mathcal{O}(\epsilon)$  is small. We also assume that  $\|h_0\|_{\nu_0} \leq C_{h_0}$  and  $\|\partial_w h_0\|_{\nu_0} \leq C'_{h_0}$ ;
- (H3) The function  $r_0 : (\mathcal{V} + i\nu_0) \times (\mathcal{S} + i\sigma_0) \rightarrow \mathbb{C}$ , is real analytic, satisfies  $r_0 = \mathcal{O}(z_0^3)$ , and its norm is uniformly bounded by  $C_{r_0}$ ;
- (H4) There exists a locally unique equilibrium  $(w^e, \zeta(w^e))$  in the  $(w, z)$ -coordinates which in the transformed  $(w_0, z_0)$ -coordinates takes the form  $(w_0^e, 0) \in \mathcal{V} \times \mathcal{S}$ .

Note that we can assume  $z_0^e = 0$  in (H4) without loss of generality. Indeed, the equilibrium  $(w^e, \zeta(w^e))$  transforms to  $(w_0^e, z_0^e)$  with

$$z_0^e = \zeta(w^e) - \zeta(u^e, v_0^e),$$

and so by introducing the affine symplectic transformation  $(w_0, z_0) \mapsto (w_0, z_0 - z_0^e)$  we move this equilibrium to  $(w_0^e, 0)$ . It is important to start our iteration from  $z_0^e = 0$  - we can then ensure that our improved slow manifolds, which we will define iteratively, contain this equilibrium. Obviously we could also transform  $w_0^e = 0$  but this will not be necessary.

## 3 Main results

### 3.1 General slow-fast system

Our main result for general slow-fast systems is the following:

**Theorem 3.1** *Consider (1.2) and assume that assumptions (G1), (G2) and (G3) hold true and let  $\underline{\nu}, \underline{\sigma}, \xi_0$  be chosen positive constants satisfying*

$$\underline{\nu} < \nu_1 < \nu_0 \quad \text{and} \quad \underline{\sigma} < \sigma_1 < \sigma_0 \quad \text{where} \quad \nu_1 = \nu_0 - \xi_0, \quad \sigma_1 = \sigma_0 - \xi_0. \quad (3.1)$$

*Then there exists an  $\bar{\epsilon} > 0$  so that for every  $\epsilon \leq \bar{\epsilon}$  the following holds true: There exist*

*(i) a transformation  $(\mathcal{V} + i\underline{\nu}) \times (\mathcal{S} + i\underline{\sigma}) \rightarrow (\mathcal{V} + i\nu_0) \times (\mathcal{S} + i\sigma_0)$ ,*

$$(w, z_*) \mapsto (w_0 = w, z_0 = z_* + \zeta_*(w)),$$

*transforming (1.2) into the normal form*

$$\begin{aligned} \partial_t w &= \epsilon W_*(w, z_*), \\ \partial_t z_* &= \rho_*(w) + A_*(w) z_* + R_*(w, z_*), \end{aligned}$$

*where  $W_*(w, z_*) = W_0(w, z_0)$ ;*

(ii) constants  $C_1, \dots, C_5$  only depending upon the previous constants so that

$$\|\rho_*(w)\| \leq C_1 \epsilon^2 \|W_*(w, 0)\| e^{-[m \ln 2 / (4K_1 \epsilon)]}, \quad m = \min\{\nu_1 - \underline{\nu}, \sigma_1 - \underline{\sigma}\}, \quad (3.2)$$

where  $K_1 = K_0 \left(1 - K_0 \underline{\sigma}^{-2} \delta_0 \left(\frac{\epsilon K_0}{\xi_0} \underline{\sigma} + 2C_{R_0} K_0\right)\right)^{-1}$ , and

$$\begin{aligned} \|z_* - z_0\|_{\underline{\nu}, \underline{\sigma}} &\leq C_2 \epsilon, \quad \|A_* - A_0\|_{\underline{\nu}} \leq C_3 \epsilon, \\ \|R_* - R_0\|_{\underline{\nu}, \underline{\sigma}} &\leq C_4 \epsilon, \quad \|W_* - W_0\|_{\underline{\nu}, \underline{\sigma}} \leq C_5 \epsilon. \end{aligned}$$

In other words: (1.2) has a slow manifold given by  $z_* = 0$  which is invariant up to the exponential small error  $\rho_*$ , which vanishes at equilibria where  $W_*(w, 0) = 0$ .  $\square$

### 3.2 Hamiltonian slow-fast system

In the Hamiltonian case it is a bit more difficult to be explicit about the dependency of new constants on the old constants without introducing an overwhelming amount of clutter. Theorem 3.2 will therefore appear slightly less explicit than Theorem 3.1.

**Theorem 3.2** Consider the Hamiltonian  $H = H(w_0, z_0)$  from (2.8) and assume that assumptions (H1), (H2) and (H3) hold true. Furthermore, let  $\underline{\nu}$ ,  $\underline{\sigma}$ ,  $\xi_0$  be chosen positive constants satisfying  $\underline{\nu} < \nu_1 < \nu_0$  and  $\underline{\sigma} < \sigma_1 < \sigma_0$  where  $\nu_1 = \nu_0 - \xi_0$  and  $\sigma_1 = \sigma_0 - \xi_0$ . Then there exists an  $\bar{\epsilon} > 0$  so that for every  $\epsilon \leq \bar{\epsilon}$  the following holds true: There exist

(i) a transformation  $(\mathcal{V} + i\underline{\nu}) \times (\mathcal{S} + i\underline{\sigma}) \rightarrow (\mathcal{V} + i\nu_0) \times (\mathcal{S} + i\sigma_0)$ ,  $(w_*, z_*) \mapsto (w_0, z_0)$  transforming (2.8) into the normal form

$$H_*(w_*, z_*) = H(w_0, z_0) = h_*(w_*) + \langle \rho_*(w_*), z_* \rangle + \frac{1}{2} \langle A_*(w_*) z_*, z_* \rangle + r_*(z_*, w_*). \quad (3.3)$$

(ii) constants  $C_1, \dots, C_5$  only depending upon the previous constants so that

$$\|\rho_*\| \leq C_1 \epsilon^2 e^{-C_2/\epsilon},$$

and

$$\begin{aligned} \|z_* - z_0\|_{\underline{\nu}, \underline{\sigma}}, \|w_* - w_0\|_{\underline{\nu}, \underline{\sigma}} &\leq C_3 \epsilon, \\ \|A_* - A_0\|_{\underline{\nu}, \underline{\sigma}}, \|r_* - r_0\|_{\underline{\nu}, \underline{\sigma}} &\leq C_5 \epsilon. \end{aligned}$$

(iii) If (H4) holds, so that  $(w_0^e, z_0^e = 0)$  is an equilibrium, then  $\rho_*|_{(w_0^e, 0)} = 0$ .

In other words: (2.8) has a slow manifold given by  $z_* = 0$  which is invariant up to the exponential small error  $\rho_*$ , and by (iii) contains the nearby equilibrium  $(w_0^e, z_0^e = 0)$ .  $\square$

We can now directly address stability:

**Corollary 3.1** Consider the Hamiltonian  $H_* = H_*(w, z)$  from (3.3) on the real domain  $\mathcal{V} \times \mathcal{S}$ . If  $A_0$  is positive definite then  $L(w, z) = \frac{1}{2}\langle A_*(w)z, z \rangle + r_*(w, z)$  is an approximate Lyapunov function for  $\|z\|$  and  $\epsilon$  sufficiently small, and there exist constants  $c_1$  and  $c_2$  so that

$$\|z(t)\| \leq \mathcal{O}(e^{-c_1/\epsilon}) \quad \text{for } 0 \leq t \leq c_2\epsilon^{-2},$$

when  $z(0) = 0$ .

PROOF By assumption  $A_*$  is positive definite for  $\epsilon$  sufficiently small, and hence there exist constants  $\underline{\lambda} > 0$  and  $\bar{\lambda} > 0$  so that

$$0 < \underline{\lambda}\|z\|^2 \leq L \leq \bar{\lambda}\|z\|^2 \quad (3.4)$$

for  $\|z\|$  small and  $w \in \mathcal{V} + i\nu$ . Let  $J_z, J_w$  be such that the symplectic form  $\omega$  from (2.6) satisfies

$$\omega((z_1, w_1), (z_2, w_2)) = \langle z_1, J_z^{-1}z_2 \rangle + \epsilon^{-1}\langle w_1, J_w^{-1}w_2 \rangle. \quad (3.5)$$

When differentiating  $L$  we then obtain,

$$\begin{aligned} \dot{L} &= \langle \partial_z L, J_z(\rho_* + A_*z + \partial_z r_*) \rangle + \epsilon \langle \partial_w L, J_w \partial_w H \rangle \\ &= \langle A_*z + \partial_z r_*, J_z(\rho_* + A_*z + \partial_z r_*) \rangle + \epsilon \langle \partial_w L, J_w \partial_w H \rangle \\ &= \langle A_*z + \partial_z r_*, J_z \rho_* \rangle + \epsilon \langle \partial_w L, J_w \partial_w h_*(w_*) \rangle \\ &\leq C_6 \epsilon^2 e^{-C_2/\epsilon} + C_7 \epsilon \sup_{w \in \mathcal{V} + i\nu} L(w, z(t)), \end{aligned}$$

on  $(w, z) \in \mathcal{V} \times \mathcal{S}$  for some constants  $C_6$  and  $C_7$ . Here we have used a Cauchy estimate on  $\partial_w L$ . Integrating this inequality from  $t = 0$  to  $t$  and using (3.4) we find that

$$\underline{\lambda}\|z\|^2 \leq L \leq C_6 \epsilon^2 t e^{-C_2/\epsilon} + C_7 \epsilon \int_0^t \sup_w L(w, z(s)) ds \leq C_6 \epsilon^2 t e^{-C_2/\epsilon} + C_7 \bar{\lambda} \epsilon \int_0^t \|z(s)\|^2 ds.$$

We have here used that  $L(0) = 0$  since  $z(0) = 0$ . Then by Gronwall's inequality in integral form [1] we obtain

$$\|z\|^2 \leq C_6 \underline{\lambda}^{-1} \epsilon^2 t e^{-C_2/\epsilon} e^{C_7 \bar{\lambda} \underline{\lambda}^{-1} \epsilon t},$$

and therefore while  $0 \leq t \leq C_2 \underline{\lambda} / (2C_7 \bar{\lambda} \epsilon^2)$ :

$$\|z\| \leq \sqrt{\frac{C_6 C_2}{2C_7 \bar{\lambda}}} e^{-C_2/(4\epsilon)},$$

completing the proof. ■

Note that this upper estimate  $\mathcal{O}(\epsilon^{-2})$  is large, even on the fast time scale  $\tau$  (1.4) where it is  $\mathcal{O}(\epsilon^{-1})$ .

## 4 Proof of Theorem 3.1

We will first state and prove two key lemmata. The first lemma will be an application of the implicit function theorem, while the second one will be “The Iterative Lemma”. We will use these lemmata successively, so assume that  $\rho$ ,  $A$  and  $R$  (in place of  $\rho_0$ ,  $A_0$  and  $R_0$ ) satisfy the assumptions (G1), (G2), and (G3).

**Lemma 4.1** *Suppose that  $0 < \kappa \leq \sigma - K\delta$  where*

$$\delta = \|\rho(w)\|_\nu < \kappa^2 K^{-2} C_R^{-1}. \quad (4.1)$$

*Then the equation*

$$0 = \rho(w) + A(w)z + R(w, z), \quad (4.2)$$

*has a locally unique solution  $z = \zeta(w) \in \mathcal{Z}_\mathbb{C}$  satisfying:*

$$\|\zeta(w)\| \leq K\|\rho(w)\|, \quad (4.3)$$

*for every  $w \in \mathcal{V} + i\nu$ . Moreover  $\zeta(w)$  is analytic in  $w \in \mathcal{V} + i\nu$ .  $\square$*

PROOF Re-arranging (4.2) and applying the inverse  $A(w)^{-1}$  gives

$$\zeta = -A(w)^{-1}(\rho(w) + R(w, \zeta)). \quad (4.4)$$

Put  $\zeta_0(w) = -A(w)^{-1}\rho(w)$  and consider  $\zeta = \zeta_0 + z$  so that

$$z = -A(w)^{-1}R(w, \zeta_0 + z). \quad (4.5)$$

Note that  $\|\zeta_0\| \leq \frac{K}{2}\|\rho(w)\|$ . We will denote this upper bound by  $r := \frac{K}{2}\|\rho(w)\|$  and highlight  $r \leq \frac{K}{2}\delta$ . Set

$$T(z) = -A(w)^{-1}R(w, \zeta_0 + z).$$

We will show that  $T$  is a contraction on  $B_r^\mathcal{Z}(0) \subset \mathcal{S} + i\sigma$ . By Taylor’s formula it follows that

$$R(w, z) = \int_0^1 (1-s) \partial_z^2 R(w, sz) z^2 ds,$$

and therefore by applying a Cauchy estimate we obtain for  $\|z\| \leq K\delta$

$$\|R(w, z)\| \leq \frac{1}{2} \sup_{0 \leq s \leq 1} \|\partial_z^2 R(w, sz)\| \|z\|^2 \leq \frac{C_R}{\kappa^2} \|z\|^2.$$

Here  $0 < \kappa \leq \sigma - K\delta$ . Therefore

$$\|T(z)\| \leq \frac{K}{2} \frac{C_R}{\kappa^2} (2r)^2 \leq K^2 \frac{C_R}{\kappa^2} \delta r < r$$

using that  $2r \leq K\delta$  and the assumption (4.1), and hence  $T : B_r^\mathcal{Z}(0) \rightarrow B_r^\mathcal{Z}(0)$ . Next, we use Taylor’s formula to obtain

$$\partial_z R(w, z) = \int_0^1 \partial_z^2 R(w, tz) z dt,$$

and then apply a Cauchy estimate to bound  $\partial_z R$ :

$$\|\partial_z R(w, \zeta_0 + z)\| \leq \frac{2C_R}{\kappa^2} \|\zeta_0 + z\|.$$

Therefore

$$\|\partial_z T(z)\| \leq K^2 \frac{C_R}{\kappa^2} \delta < 1, \quad \blacksquare$$

using (4.1). This shows that  $T$  is a contraction on  $B_r^{\mathcal{Z}}(0)$  and there exists a unique fix point  $z(w)$  of  $T$ . In particular,  $\zeta(w) = \zeta_0(w) + z(w)$  solves (4.4) and  $\|\zeta(w)\| \leq 2r = K\|\rho(w)\|$ . By [10, Section 1.2.6] the map  $\zeta : \mathcal{V} + i\nu \rightarrow \mathcal{Z}$  is analytic (this is where we need  $w \rightarrow A^{-1}(w)$  to be analytic).

**Lemma 4.2** (*The Iterative Lemma*) *Let*

$$\xi \geq 2K\delta \tag{4.6a}$$

*and*

$$\delta \leq \frac{1}{2} \kappa^2 K^{-2} C_R^{-1}, \quad \epsilon < 2\kappa^2, \tag{4.6b}$$

$\nu - \xi > 0$  and  $0 < \kappa \leq \sigma - \xi$ . Then the transformation

$$(w, z_+) \mapsto (w, \zeta(w) + z_+)$$

from  $(\mathcal{V} + i(\nu - \xi)) \times (\mathcal{S} + i(\sigma - \xi))$  into  $(\mathcal{V} + i\nu) \times (\mathcal{S} + i(\sigma - \xi/2))$  transforms

$$\begin{aligned} \partial_t w &= \epsilon W(w, z), \\ \partial_t z &= \rho(w) + A(w)z + R(w, z), \end{aligned}$$

into the normal form

$$\begin{aligned} \partial_t w &= \epsilon W_+(w, z_+), \\ \partial_t z_+ &= A_+(w)z_+ + \rho_+(w) + R_+(w, z_+). \end{aligned}$$

Here  $\|\rho\|_{\chi} \leq \delta$ ,  $\zeta$  is the solution from Lemma 4.1, we define  $W_+(w, z_+) = W(w, \zeta + z_+)$  and the following estimates hold true

$$\|\rho_+(w)\| \leq \frac{\epsilon K}{\xi} \delta \|W_+(w, 0)\|, \tag{4.7a}$$

$$\|A_+(w)^{-1}\| \leq \frac{K}{2} \left( 1 - \frac{K\delta}{2} \kappa^{-2} \left( \frac{\epsilon K}{\xi} \kappa + 2C_R K \right) \right)^{-1}, \tag{4.7b}$$

$$\|A_+ - A\|_{\nu-\xi} \leq \kappa^{-2} \delta \left( \frac{\epsilon K}{\xi} \kappa + 2C_R K \right), \tag{4.7c}$$

$$\|R_+ - R\|_{\nu-\xi, \sigma-\xi} \leq \frac{\epsilon K}{\xi} \delta (2 + \kappa^{-1} C_z) + C_R \left( \frac{2K\delta}{\xi} + \kappa^{-2} K \delta (K\delta + 2C_z) \right), \tag{4.7d}$$

for  $w \in \mathcal{V} + i(\nu - \xi)$  and where  $\|A(w)^{-1}\|_{\nu} \leq \frac{K}{2}$ ,  $\|R\|_{\nu, \sigma} \leq C_R$  and  $\|z\| \leq C_z$ . In particular,  $\delta_+ = \|\rho_+\|_{\nu-\xi}$  satisfies

$$\delta_+ \leq \frac{\epsilon K}{\xi} \delta. \tag{4.8}$$

□

PROOF Applying the transformation  $z = \zeta + z_+$  gives

$$\begin{aligned}\partial_t w &= \epsilon W_+(w, z_+), \\ \partial_t z_+ &= A_+(w)z_+ + \rho_+(w) + R_+(w, z_+),\end{aligned}$$

and using the fact that  $\zeta(w)$  solves (4.2) we obtain

$$A_+(w) = (A(w) - \epsilon \partial_w \zeta(w) \partial_z W(w, \zeta(w)) + \partial_z R(w, \zeta(w))), \quad (4.9a)$$

$$\rho_+(w) = -\epsilon \partial_w \zeta(w) W(w, \zeta(w)), \quad (4.9b)$$

$$W_+(w, z_+) = W(w, \zeta + z_+), \quad (4.9c)$$

and

$$R_+(w, z_+) = -\epsilon \partial_w \zeta(w) \tilde{W}(w, z_+) + \tilde{R}(w, z_+), \quad (4.10)$$

$$\tilde{W}(w, z_+) = W(w, \zeta + z_+) - W(w, \zeta) - \partial_z W(w, \zeta) z_+, \quad (4.11)$$

$$\tilde{R}(w, z_+) = R(w, \zeta + z_+) - R(w, \zeta) - \partial_z R(w, \zeta) z_+. \quad (4.12)$$

We have

$$\|\rho_+(w)\| \leq \epsilon \frac{\|\zeta\|_\nu}{\xi} \|W(w, \zeta)\| \leq \frac{\epsilon K}{\xi} \delta \|W_+(w, 0)\|, \quad (4.13)$$

for  $w \in \mathcal{V} + i(\nu - \xi)$ . Here we have used the Cauchy estimate:

$$\|\partial_w \zeta\|_{\nu-\xi} \leq \frac{\|\zeta\|_\nu}{\xi}, \quad (4.14)$$

for  $\nu - \xi > 0$  and  $\xi > 0$  and Lemma 4.1. To complete one step of the iteration we also have to estimate

$$A_+(w) - A(w) = -\epsilon \partial_w \zeta \partial_z W(w, \zeta) + \partial_z R(w, \zeta),$$

and  $R_+ - R$  appropriately. For the former first note that by (4.14) for  $w \in \mathcal{V} + i(\nu - \xi)$

$$\begin{aligned}\|\partial_w \zeta \partial_z W(w, \zeta)\|_{\mathcal{L}(\mathcal{Z}, \mathcal{Z})} &\leq \|\partial_w \zeta\|_{\mathcal{L}(\mathcal{W}, \mathcal{Z})} \|\partial_z W(w, \zeta)\|_{\mathcal{L}(\mathcal{Z}, \mathcal{W})} \\ &\leq \frac{\|\zeta\|_\nu}{\xi} \kappa^{-1}.\end{aligned}$$

using another Cauchy estimate, noting (2.5) and  $\sigma - \xi \geq \kappa > 0$  and  $\|\zeta\|_\nu \leq \xi/2$  by Lemma 4.1 and (4.6a). Moreover

$$\|\partial_z R(w, \zeta)\| = \left\| \int_0^1 \partial_z^2 R(w, s\zeta) s \zeta ds \right\| \leq 2C_R \kappa^{-2} \|\zeta\|_\nu,$$

for all  $w \in \mathcal{V} + i(\nu - \xi)$ , and therefore for all such  $w$ , as  $\|\zeta\|_\nu \leq K\delta$  by Lemma 4.1,

$$\|A_+ - A\|_{\nu-\xi} \leq \frac{\epsilon K}{\xi} \delta \kappa^{-1} + 2C_R \kappa^{-2} K \delta = \kappa^{-2} \delta \left( \frac{\epsilon K}{\xi} \kappa + 2C_R K \right).$$

Hence,

$$\begin{aligned}
\frac{2}{K}\|z\| &\leq \|A(w)z\| \\
&= \|(A_+(w) - (A_+(w) - A(w))z)\| \\
&\leq \|A_+(w)z\| + \|A_+(w) - A(w)z\| \\
&\leq \|A_+(w)z\| + \|A_+(w) - A(w)\|_{\nu-\xi}\|z\| \\
&\leq \|A_+(w)z\| + \kappa^{-2}\delta \left( \frac{\epsilon K}{\xi}\kappa + 2C_R K \right) \|z\|
\end{aligned}$$

so that

$$\frac{2}{K} \left( 1 - \frac{K\kappa^{-2}}{2}\delta \left( \frac{\epsilon K}{\xi}\kappa + 2C_R K \right) \right) \|z\| \leq \|A_+(w)z\|,$$

if

$$\frac{K\kappa^{-2}}{2}\delta \left( \frac{\epsilon K}{\xi}\kappa + 2C_R K \right) < 1.$$

This is true due to (4.6):

$$\begin{aligned}
\frac{K\kappa^{-2}}{2}\delta \left( \frac{\epsilon K}{\xi}\kappa + 2C_R K \right) &\leq \frac{\kappa^{-2}\epsilon}{4} + K^2\kappa^{-2}C_R\delta \\
&< \frac{1}{2} + \frac{1}{2} = 1,
\end{aligned}$$

We highlight that by (4.6b) this choice of  $\delta$  satisfies the condition (4.1) required in Lemma 4.2 - it is in fact a factor 2 stronger. Hence

$$\|A_+(w)^{-1}\| \leq \frac{K}{2} \left( 1 - \frac{K^2\kappa^{-2}\delta}{2}(\epsilon\kappa\xi^{-1} + 2C_R) \right)^{-1},$$

for all  $w \in \mathcal{V} + i(\nu - \xi)$ .

For  $R_+ - R$ , with  $R_+$  from (4.10) we first estimate  $\tilde{W}$  from (4.11) and  $\tilde{R} - R$  from (4.12). This gives:

$$\|\tilde{W}\|_{\nu, \sigma-\xi} \leq 2 + \kappa^{-1}C_z$$



and

$$\begin{aligned}
\|\tilde{R} - R\|_{\nu-\xi, \sigma-\xi} &\leq \|R(w, \zeta + z_+) - R(w, z_+)\|_{\nu-\xi, \sigma-\xi} + \|R(w, \zeta)\|_{\nu-\xi, \sigma-\xi} \\
&\quad + \|\partial_z R(w, \zeta) z_+\|_{\nu-\xi, \sigma-\xi} \\
&\leq \left\| \int_0^1 \partial_z R(w, s\zeta + z_+) ds \zeta \right\|_{\nu-\xi, \sigma-\xi} + \left\| \int_0^1 (1-t) \partial_z^2 R(w, t\zeta) \zeta^2 dt \right\|_{\nu-\xi, \sigma-\xi} \\
&\quad + \left\| \int_0^1 \partial_z^2 R(w, t\zeta) \zeta z_+ dt \right\|_{\nu-\xi, \sigma-\xi} \\
&\leq \|\partial_z R\|_{\nu, \sigma-\xi/2} \|\zeta\|_\nu + \frac{1}{2} \frac{2C_R}{\kappa^2} \|\zeta\|_\nu^2 + \frac{2C_R}{\kappa^2} \|\zeta\|_\nu C_z \\
&\leq \frac{2C_R}{\xi} K\delta + C_R \kappa^{-2} K\delta (K\delta + 2\sigma) \\
&= C_R \left( \frac{2K\delta}{\xi} + \kappa^{-2} K\delta (K\delta + 2\sigma) \right),
\end{aligned}$$

Therefore

$$\begin{aligned}
\|R_+ - R\|_{\nu-\xi, \sigma-\xi} &\leq \epsilon \|\partial_w \zeta\|_{\nu-\xi} \|\tilde{W}\|_{\nu-\xi, \sigma-\xi} + \|\tilde{R} - R\|_{\nu-\xi, \sigma-\xi} \\
&\leq \frac{\epsilon K}{\xi} \delta (2 + \kappa^{-1} C_z) + C_R \left( \frac{2K\delta}{\xi} + \kappa^{-2} K\delta (K\delta + 2C_z) \right). \quad \blacksquare
\end{aligned}$$

Following this lemma, we introduce new constants:

$$K_+ = K \left( 1 - \frac{K\kappa^{-2}\delta}{2} \left( \frac{\epsilon K}{\xi} \kappa + 2C_R K \right) \right)^{-1}, \quad (4.15a)$$

$$C_{R_+} = C_R \left( 1 + \frac{2K\delta}{\xi} + \kappa^{-2} K\delta (K\delta + 2C_z) \right) + \frac{\epsilon K\delta}{\xi} (2 + \kappa^{-1} C_z), \quad (4.15b)$$

$$\sigma_+ = \sigma - \xi,$$

$$\nu_+ = \nu - \xi.$$

We will now apply Lemma 4.2 to (1.2) once with  $\xi = \xi_0 = \mathcal{O}(1) < \max(\nu_0 - \underline{\nu}, \sigma_0 - \underline{\sigma})$  large. Due to (2.3) and (4.7a) there is a  $C_\delta > 0$  such that

$$\delta_1 = C_\delta \epsilon^2. \quad (4.16)$$

So we start from  $(\mathcal{V} + i\nu_1) \times (\mathcal{S} + i\sigma_1)$  with  $\nu_1 = \nu_0 - \xi_0$  and  $\sigma_1 = \sigma_0 - \xi_0$ . By (2.3) and (4.3) we have  $\|z_1 - z_0\|_\nu \leq K_0 \delta$ . Moreover we can bound  $A_1^{-1}$  and  $R_1$  by

$$\|A_1^{-1}\|_{\nu_1} \leq K_1 = K_0 \left( 1 - \frac{K_0 \underline{\sigma}^{-2} \delta_0}{2} \left( \frac{\epsilon K_0}{\xi_0} \underline{\sigma} + 2C_{R_0} K_0 \right) \right)^{-1},$$

resp.

$$\|R_1\|_{\nu_1, \sigma_1} \leq C_{R_1} = C_{R_0} \left( 1 + \frac{2K_0 \delta_0}{\xi_0} + \underline{\sigma}^{-2} K_0 \delta_0 (K_0 \delta_0 + 2C_z) \right) + \frac{\epsilon K_0 \delta_0}{\xi_0} (2 + \underline{\sigma}^{-1} C_z),$$

taking  $\kappa = \underline{\sigma}$  in (4.15a) and (4.15b) (noting that by (3.1) this choice of  $\kappa$  satisfies  $\kappa \leq \sigma - \xi_0 = \sigma_1$  as required in Lemma 4.2). We then continue to apply Lemma 4.2 with

$$\xi_n = 2K_n\epsilon, \quad (4.17)$$

for  $n \leq N$ , where  $N$  is such that  $\nu_N - \xi_N \geq \underline{\nu}$  and  $\sigma_N - \xi_N \geq \underline{\sigma}$ . Here we have defined

$$\begin{aligned} \nu_{n+1} &= \nu_1 - \sum_{i=1}^n \xi_i, \\ \sigma_{n+1} &= \sigma_1 - \sum_{i=1}^n \xi_i. \end{aligned}$$

Note that the form of  $\xi_n$  from (4.17) allows us to apply “The Iterative Lemma 4.2”, since at each step

$$\xi_n \geq 2K_n\delta_1 = 2K_nC_\delta\epsilon^2 \geq 2K_n\delta_n$$

provided  $\epsilon$  is sufficiently small so that  $\epsilon \leq C_\delta^{-1}$  and provided that  $\delta_n \leq \delta_1$ . Here we used (4.16). Then by (4.8),

$$\begin{aligned} \delta_{n+1} &\leq \epsilon \frac{K_n\delta_n}{\xi_n} \leq \epsilon^n \frac{K_n \times \cdots \times K_1}{\xi_n \times \cdots \times \xi_1} \delta_1 \\ &\leq 2^{-n}\delta_1. \end{aligned} \quad (4.18)$$

Here  $\delta_{n+1} = \|\rho_{n+1}\|_{\nu_{n+1}}$ . To prove the theorem, it is left to be shown that  $\xi_n = O(\epsilon)$  with an order constant which is bounded in  $n$  so that by choosing  $N = \mathcal{O}(\epsilon^{-1})$  iteration steps the estimates in Theorem 3.1 apply. We therefore estimate  $K_{n+1}$  and  $C_{R_{n+1}}$ :

$$\begin{aligned} K_{n+1} &\leq K_n \left( 1 - \frac{K_n}{2} \underline{\sigma}^{-2} \delta_n \left( \frac{\epsilon K_n}{\xi} \kappa + 2C_{R_n} K_n \right) \right)^{-1} \\ &\leq K_n \left( 1 - 2^{-n+1} K_n \underline{\sigma}^{-2} \delta_1 \left( \frac{1}{2} \underline{\sigma} + 2C_{R_n} K_n \right) \right)^{-1}, \\ C_{R_{n+1}} &\leq C_{R_n} \left( 1 + \frac{2K_n\delta_n}{\xi_n} + \underline{\sigma}^{-2} K_n \delta_n (K_n \delta_n + 2C_z) \right) + \frac{\epsilon K_n \delta_n}{\xi_n} (2 + \kappa^{-1} C_z) \\ &\leq C_{R_n} (1 + 2^{-n+1} \epsilon^{-1} \delta_1 + 2^{-n} \underline{\sigma}^{-2} K_n \delta_1 (2^{-n} K_n \delta_1 + 2C_z)) + 2^{-n} \delta_1 (2 + \underline{\sigma}^{-1} C_z). \end{aligned} \quad (4.19)$$

If

$$\underline{\sigma}^{-2} \delta_1 \left( \frac{1}{2} \underline{\sigma} + 2C_{R_n} K_n \right) \leq \frac{1}{8K_1} \quad (4.21)$$

and  $\alpha_i = \frac{K_i}{K_1}$  then (4.19) implies

$$\alpha_{n+1} \leq (1 - 2^{-n-2} \alpha_n)^{-1} \alpha_n. \quad (4.22)$$

It is easy to show that

$$\alpha_n \leq \frac{2^{n+1}}{2^n + 1} \leq 2, \quad (4.23)$$

and so

$$K_n \leq 2K_1 \quad \text{for all } n \leq N. \quad (4.24)$$

This estimate could be improved, but we do not need to do so. Similarly for  $C_{R_n}$ , we set  $\beta_n = \frac{C_{R_n}}{C_{R_1}}$  and obtain

$$\beta_{n+1} \leq \beta_n (1 + 2^{-n+1}\epsilon^{-1}\delta_1 + 2^{-n+2}\underline{\sigma}^{-2}K_1\delta_1 (2^{-n}K_1\delta_1 + C_z)) + 2^{-n}\tilde{\delta}_1$$

with  $\tilde{\delta}_1 = C_{R_1}^{-1}\delta_1(2 + \underline{\sigma}^{-1}C_z)$ , using the estimate (4.24) obtained for  $K_n$ . Let

$$\underline{\sigma}^{-2}K_1\epsilon(K_1\delta_1 + C_z) \leq \frac{1}{2} \quad (4.25)$$

so that

$$\beta_{n+1} \leq (1 + 2^{-n+2}C_\delta\epsilon)\beta_n + 2^{-n}\tilde{\delta}_1. \quad (4.26)$$

Let  $N \in \mathbb{N}$  be such that  $\beta_n \leq 2$  for  $n \leq N$ . Then for  $n \leq N$  we get, using the geometric series, that

$$\beta_n \leq \beta_1 + 8C_\delta\epsilon + \tilde{\delta}_1. \quad (4.27)$$

Using (4.13) and choosing  $\epsilon$  small we see that  $\beta_n \leq 2$  for all  $n$  or simply

$$C_{R_n} \leq 2C_{R_1}. \quad (4.28)$$

Returning to our assumptions, (4.21) and (4.25), we realize that due to (4.16) it suffices to take  $\epsilon$  so that the inequalities

$$\begin{aligned} \underline{\sigma}^{-2}C_\delta\epsilon^2 \left( \frac{1}{2}\underline{\sigma} + 8C_{R_1}K_1 \right) &\leq \frac{1}{8K_1}, \\ \underline{\sigma}^{-2}K_1\epsilon(K_1C_\delta\epsilon^2 + C_z) &\leq \frac{1}{4}, \end{aligned}$$

are satisfied. Notice also that the hypotheses of Lemma 4.1 are satisfied at each step. Therefore, by (4.17) and (4.24),

$$\xi_n \leq 4K_1\epsilon.$$

and so we can perform

$$N = \left\lceil \frac{m}{4K_1\epsilon} \right\rceil, \quad m = \min\{\sigma_1 - \underline{\sigma}, \nu_1 - \underline{\nu}\}.$$

iterations before one of the conditions  $\nu_{N+1} > \underline{\nu}$ ,  $\sigma_{N+1} > \underline{\sigma}$  is violated.

If we inspect  $\rho_{N+1}(w)$  further we realize that the error field estimate can be made *point-wise* in  $x$  improving near actual equilibria: by (4.13) and (4.18) we have

$$\begin{aligned} \|\rho_{N+1}(w)\| &\leq \frac{\epsilon K_N}{\xi_N} \delta_N |W_N(w, 0)| \\ &\leq 2^{-[m/(4K_1\epsilon)]} \delta_1 |W_N(w, 0)| \\ &\leq C_\delta \epsilon^2 e^{-[m \ln 2 / (4K_1\epsilon)]} |W_N(w, 0)|. \end{aligned}$$

Finally, provided  $\epsilon$  is sufficiently small, there exists constants  $c_1, \dots, c_4$  depending only upon the previous constants so that by (4.3), (4.18), (4.24) and (4.16)

$$\|z_{N+1} - z_1\|_{\mathcal{L}} \leq \sum_{j=1}^N \|\zeta_j\|_{\mathcal{L}} \leq \sum_{j=1}^N 2^{-j} K_1 \delta_1 \leq c_1 \epsilon^2. \quad (4.29)$$

Similarly, by (4.7c), using (4.17) and that  $K_n$  and  $C_{R_n}$  are bounded independently from  $n$  (see (4.24) and (4.28)), (4.18), and (4.16) we get

$$\|A_{N+1} - A_1\|_{\mathcal{L}} \leq c_2 \epsilon^2.$$

Moreover, using (4.7d), (4.17), (4.24), (4.28), (4.18), and (4.16) for the first estimate and the mean value theorem and (4.29) for the second estimate we get

$$\|R_{N+1} - R_1\|_{\mathcal{L}, \mathcal{G}} \leq c_3 \epsilon^2, \quad \|W_{N+1} - W_1\|_{\mathcal{L}, \mathcal{G}} \leq c_4 \epsilon^2.$$

Since  $\|z_1 - z_0\|_{\mathcal{L}}$  is  $\mathcal{O}(\epsilon)$  by Lemma 4.1 and (G2) the theorem follows.

**Remark 4.1** Again we highlight (cf. Example 1.1) that it is in general not possible to obtain an estimate of the form  $\mathcal{O}(e^{-C/(\epsilon \|W_*(w,0)\|)})$ , since at each step of our iteration we must apply the Cauchy estimate (4.14) forcing us to estimate  $\zeta$  on  $\mathcal{V} + i\nu$  - through (4.3) this gives rise to the factor  $C_{W_0} = 1$  rather than  $\|W_i(w, 0)\|$  at each step.  $\square$

## 5 Proof of Theorem 3.2

The proof of Theorem 3.2 is again based on an iterative lemma. As we will use this lemma successively we will in the following consider *the normal form* Hamiltonian

$$H(w, z) = h(w) + \langle \rho(w), z \rangle + \frac{1}{2} \langle A(w)z, z \rangle + r(w, z), \quad (5.1)$$

(in place of  $H_0$  from (2.8)) satisfying the assumptions (H1), (H2), (H3) and (H4) on  $(w, z) \in (\mathcal{V} + i\nu) \times (\mathcal{S} + i\sigma)$ , with  $\delta = \|\rho(w)\|_{\nu}$  small. Let  $z = \zeta(w)$  be the solution of

$$\partial_z H(w, z) = \rho(w) + A(w)z + \partial_z r(w, z).$$

If  $\|A(w)^{-1}\| \leq \frac{K}{2}$  then by Lemma 4.1 this solution satisfies the estimate  $\|\zeta\|_{\nu} \leq K\delta$  for

$$\delta < \frac{1}{3} \kappa^3 K^{-2} C_r^{-1}, \quad (5.2)$$

where the “extra” factor  $\frac{1}{3}\kappa$  compared to above in Lemma 4.1 comes from estimating  $\partial_z r = \mathcal{O}(z^2)$  rather than just  $R$ . Let  $\zeta(w) = (\zeta^x(w), \zeta^y(w))$  and consider the generating function  $G(u, v_+, x, y_+)$  from (2.7). We then consider the symplectic transformation  $(w_+, z_+) \mapsto (w, z)$  defined implicitly by the equations

$$\begin{aligned} x_+ &= \partial_{y_+} G = x - \zeta^x(u, v_+), & y &= \partial_x G = y_+ + \zeta^y(u, v_+), \\ u_+ &= \epsilon \partial_{v_+} G = u + \epsilon \partial_{v_+} g(u, v_+, x, y_+), & v &= \epsilon \partial_u G = v_+ + \epsilon \partial_u g(u, v_+, x, y_+). \end{aligned}$$

**Lemma 5.1** *The function  $g$  from (2.7) satisfies  $\|g\|_{\nu,\sigma} \leq C_D \|\zeta\|_\nu$ , where  $C_D$  only depends on the domain  $(\mathcal{V} + i\nu) \times (\mathcal{S} + i\sigma)$ .  $\square$*

PROOF This is straightforward and follows directly from the estimate

$$\|g\| \leq \|\zeta^x\| \|y_+\| + \|\zeta^y\| \|x\| \leq C_D \|\zeta\|. \quad \blacksquare$$

We will write this transformation as

$$z = z_+ + \alpha(w_+, z_+) = z_+ + \zeta(\hat{w}), \quad w = w_+ + \epsilon\beta(w_+, z_+), \quad (5.3a)$$

with  $z_+ = (x_+, y_+)$ ,  $w_+ = (u_+, v_+)$ ,  $\hat{w} = (u, v_+)$ , and use

$$\beta = \beta_0(w_+) + \langle \beta_1(w_+), z_+ \rangle + \frac{1}{2} \langle \beta_2(w_+, z_+) z_+, z_+ \rangle, \quad (5.3b)$$

and

$$\tilde{w} = w_+ + \epsilon(\tilde{\beta}_0 + \langle \tilde{\beta}_1, z_+ \rangle + \frac{1}{2} \langle \tilde{\beta}_2 z_+, z_+ \rangle). \quad (5.3c)$$

Here  $\tilde{\beta}_i = (\beta_i^u, 0)$  when  $\beta_i = (\beta_i^u, \beta_i^v)$ .

We have

**Lemma 5.2** *Let  $g$  satisfy  $\|g\|_{\nu,\sigma} \leq \frac{\xi^2}{8}$ , and let  $\xi$  be such that  $\nu - \xi > 0$  and  $\sigma - \xi > 0$ . Then  $(w_+, z_+) \mapsto (w, z)$  is a symplectic transformation  $(\mathcal{V} + i(\nu - \xi)) \times (\mathcal{S} + i(\sigma - \xi)) \rightarrow (\mathcal{V} + i(\nu - \xi/2)) \times (\mathcal{S} + i(\sigma - \xi/2))$  satisfying*

$$\|\alpha\|_{\nu-\xi,\sigma-\xi}, \|\beta\|_{\nu-\xi,\sigma-\xi} \leq \frac{4\|g\|_{\nu,\sigma}}{\xi} \leq \frac{\xi}{2}, \quad (5.4) \quad \square$$

PROOF This is basically Lemma 2 in [8]. For comparison note that their  $\epsilon s$ ,  $\epsilon \alpha$ ,  $\epsilon \beta$  is our  $g$ ,  $\alpha$  resp.  $\beta$ .  $\blacksquare$

In particular

$$\epsilon \|\beta_0\|_{\nu-\xi} = \|w_+ - w(w_+, 0)\|_{\nu-\xi} \leq \frac{4\epsilon \|s\|_\nu}{\xi}, \quad (5.5a)$$

and using Cauchy-estimates on  $\partial_{z_+} \beta(w_+, 0)$  and  $\partial_{z_+}^2 \beta(w_+, 0)$  we get

$$\|\beta_1(w_+)\| \leq \kappa^{-1} \|\beta\|_{\nu-\xi}, \quad \|\beta_2(w_+, 0)\| \leq 2\kappa^{-2} \|\beta\|_{\nu-\xi}. \quad (5.5b)$$

**Corollary 5.1** *The estimate for  $\|\alpha\|_{\nu-\xi,\sigma-\xi}$  in (5.4) can be improved to*

$$\|\alpha\|_{\nu-\xi,\sigma-\xi} \leq \|\zeta\|_\nu.$$

PROOF By the previous lemma  $\tilde{w}|_{\nu-\xi,\sigma-\xi} \subset \mathcal{V} + i(\nu - \xi/2)$  and so

$$\|\alpha\|_{\nu-\xi,\sigma-\xi} = \|\zeta(\tilde{w})\|_{\nu-\xi,\sigma-\xi} \leq \|\zeta\|_{\nu-\xi/2} \leq \|\zeta\|_\nu. \quad \blacksquare$$

The fact that  $(w_+, z_+) \mapsto (w, z)$  is well-defined with domain  $(\mathcal{V} + i(\nu - \xi)) \times (\mathcal{S} + i(\sigma - \xi))$  and co-domain  $(\mathcal{V} + i(\nu - \xi/2)) \times (\mathcal{S} + i(\sigma - \xi/2))$  was crucial here and will be in the following. The  $\nu - \xi/2$  and  $\sigma - \xi/2$  terms in the co-domains allow for a step of  $\xi/2$  to apply Lemma 2.1 to estimate derivatives on  $(\mathcal{V} + i(\nu - \xi/2)) \times (\mathcal{S} + i(\sigma - \xi/2))$  by function values on the larger domain  $(\mathcal{V} + i\nu) \times (\mathcal{S} + i\sigma)$ , c.f. the Cauchy estimate (2.1). This introduces a factor of  $2\xi^{-1}$ .

We are now ready to state and prove the “Iterative Lemma for Hamiltonian systems”:

**Lemma 5.3** (*The Iterative Lemma for Hamiltonian systems*) *Consider (5.1) and assume that the assumptions (H1), (H2) and (H3) hold true, so that  $h$ ,  $A$  and  $r$  satisfy*

$$\begin{aligned} \|h\|_\nu &\leq C_h, \|\partial_w h\|_{\nu-\xi/2} \leq C'_h, \\ \|A\|_\nu &\leq C_A, \|A^{-1}\|_\nu \leq \frac{K}{2}, \\ \|r\|_\nu &\leq C_r. \end{aligned}$$

Furthermore, let

$$\xi \geq 2K\delta, \quad KC_D\delta \leq \frac{\xi^2}{8}, \quad \delta \leq \frac{1}{3}K^{-2}C_r^{-1}\kappa^3, \quad \nu - \xi > 0 \quad \text{and} \quad 0 < \kappa \leq \sigma - \xi. \quad (5.6a)$$

Then there exists an  $\bar{\epsilon}$  and constants  $c_{h,1}$ ,  $c_{h,2}$ ,  $c_\delta$ ,  $c_{A,1}$ ,  $c_{A,2}$ ,  $c_{r,1}$  and  $c_{r,2}$ , depending only on the previous constants, so that for  $\epsilon \leq \bar{\epsilon}$  the symplectic transformation  $\Psi_+ : (w_+, z_+) \mapsto (w, z)$  from Lemma 5.2 mapping  $(\mathcal{V} + i(\nu - \xi)) \times (\mathcal{S} + i(\sigma - \xi))$  into  $(\mathcal{V} + i(\nu - \xi/2)) \times (\mathcal{S} + i(\sigma - \xi/2))$  transforms  $H = H(w, z)$  from (5.1) into

$$H_+(w_+, z_+) = H(w, z) = h_+(w_+) + \langle \rho_+(w_+), z_+ \rangle + \frac{1}{2} \langle A_+(w_+)z_+, z_+ \rangle + r_+(w_+, z_+), \quad (5.6b)$$

with

$$\begin{aligned} \|h_+ - h\|_{\nu-\xi} &\leq c_{h,1}\delta + \frac{c_{h,2}\epsilon\delta}{\xi}, \quad \|\partial_w h_+\|_{\nu-3\xi/2} \leq C'_h + \frac{2c_{h,1}\delta}{\xi} + \frac{2c_{h,2}\epsilon\delta}{\xi^2}, \quad \|\rho_+\|_{\nu-\xi} \leq \frac{c_\delta\epsilon\delta}{\xi}, \\ \|A_+ - A\|_{\nu-\xi} &\leq c_{A,1}\delta + \frac{c_{A,1}\epsilon\delta}{\xi^2}, \quad \|r_+ - r\|_{\nu-\xi, \sigma-\xi} \leq c_{r,1}\delta + \frac{c_{r,2}\epsilon\delta}{\xi^2}. \end{aligned} \quad (5.6c)$$

In particular,

$$\|(A_+)^{-1}\|_{\nu-\xi} \leq \frac{K_+}{2} := \frac{K}{2} \left( 1 - \frac{K}{2} \left( \frac{\epsilon c_{A,1}}{\xi^2} + c_{A,2} \right) \delta \right)^{-1}, \quad (5.6d)$$

and  $\delta_+ = \|\rho_+\|_{\nu-\xi}$  satisfies

$$\delta_+ \leq \frac{\epsilon c_\delta}{\xi} \delta. \quad (5.6e)$$

□

**Remark 5.1** The constants  $c_{h,1}$ ,  $c_{h,2}$ ,  $c_\delta$ ,  $c_{A,1}$ ,  $c_{A,2}$  and  $c_{r,1}$ ,  $c_{r,2}$  depend polynomially on  $K$ ,  $C_H$ ,  $C_h$ ,  $C'_h$ ,  $C_A$ ,  $C_r$  and  $\delta$ . This is made more explicit in the proof below. □

PROOF It is convenient to introduce

$$\underline{w}_+ = w|_{z_+=0} = w_+ + \epsilon\beta_0(w_+), \quad (5.7)$$

and

$$\zeta_s = \zeta(\underline{w}_+ + s\epsilon(\tilde{\beta}_0 - \beta_0)), \quad (5.8)$$

with  $s \in [0, 1]$ , so that in particular  $\zeta_0 = \zeta(\underline{w}_+)$  and  $\zeta_1 = \zeta(w_+ + \epsilon\tilde{\beta}_0)$ . Note also, due to (5.3b) and (5.3c), that

$$\zeta_1 = \zeta(\tilde{w}|_{z_+=0}). \quad (5.9)$$

We will use this frequently in the following where we expand the new Hamiltonian  $H_+(w_+, z_+) = H(w, z)$  about  $z_+ = 0$  to put it in the normal form (5.6b) with

$$h_+(w_+) = H_+(w_+, 0), \quad \rho_+(w_+) = \partial_{z_+} H_+(w_+, 0), \quad A_+(w_+) = \partial_{z_+}^2 H_+(w_+, 0),$$

and

$$r_+ = H_+ - (h_+ + \langle \rho_+, z_+ \rangle + \frac{1}{2} \langle A_+ z_+, z_+ \rangle).$$

Due to (5.7) and (5.9) we obtain

$$\begin{aligned} h_+(w_+) &= H_+(w_+, 0) = H(\underline{w}_+, \zeta_1) \\ &= H(\underline{w}_+, 0) + \int_0^1 \partial_z H(\underline{w}_+, s\zeta_1) \zeta_1 ds, \\ &= h(w_+ + \epsilon\beta_0) + \int_0^1 \partial_z H(\underline{w}_+, s\zeta_1) \zeta_1 ds, \\ &= h(w_+) + \int_0^1 (\epsilon \partial_w h(w_+ + s\epsilon\beta_0) \beta_0 + \partial_z H(\underline{w}_+, s\zeta_1) \zeta_1) ds. \end{aligned} \quad (5.10)$$

Using that by (5.6a) and Lemma 4.1 we have  $\xi \geq 2K\delta \geq 2\|\zeta_1\|_{\nu-\xi/2}$  and employing (5.5a) we obtain

$$\begin{aligned} \|h_+(w_+) - h(w_+)\|_{\nu-\xi} &\leq \epsilon \|\partial_w h\|_{\nu-\xi/2} \|\beta_0\|_{\nu-\xi} + \|\partial_z H\|_{\nu-\xi/2, \xi/2} \|\zeta\|_{\nu-\xi/2} \\ &\leq \epsilon C'_h \frac{4C_D K \delta}{\xi} + \|\partial_z H\|_{\nu-\xi/2, \xi/2} K \delta \\ &\leq C'_h \frac{4\epsilon C_D K}{\xi} \delta + \kappa^{-1} C_H K \delta \\ &\leq c_{h,1} \delta + \frac{c_{h,2} \epsilon \delta}{\xi}, \end{aligned} \quad (5.11)$$

introducing the constants

$$c_{h,1} = \kappa^{-1} C_H K, \quad c_{h,2} = 4C_D K C'_h.$$

In (5.11) we have used (5.5a), Lemma 5.1 and Lemma 4.1.



Moreover, using (5.3a) and (5.3c) to compute  $\frac{\partial \tilde{w}}{\partial z_+}|_{z_+=0}$  and  $\frac{\partial \zeta}{\partial z_+}|_{z_+=0}$  we obtain

$$\begin{aligned}\rho_+(w_+) &= \partial_{z_+} H_+(w_+, 0) = \partial_z H(\underline{w}_+, \zeta_1)(I + \epsilon \partial_w \zeta_1 \tilde{\beta}_1) + \epsilon \partial_w H(\underline{w}_+, \zeta_1) \beta_1 \\ &= \epsilon \int_0^1 \partial_z^2 H(\underline{w}_+, \zeta_s) \partial_w \zeta_s (\tilde{\beta}_0 - \beta_0) ds (I + \epsilon \partial_w \zeta \tilde{\beta}_1) + \epsilon \partial_w H(\underline{w}_+, \zeta_1) \beta_1,\end{aligned}\quad (5.12)$$

where we have used that  $\partial_z H(w, \zeta(w)) = 0$  to arrive at

$$\begin{aligned}\partial_z H(\underline{w}_+, \zeta_1) &= \partial_z H(\underline{w}_+, \zeta(\underline{w}_+ + \epsilon(\tilde{\beta}_0 - \beta_0))) = \underbrace{\partial_z H(\underline{w}_+, \zeta(\underline{w}_+))}_{=0} + \epsilon \int_0^1 \frac{d}{d\epsilon} \partial_z H(\underline{w}_+, \zeta_s(\epsilon)) ds \\ &= \epsilon \int_0^1 \partial_z^2 H(\underline{w}_+, \zeta_s) \partial_w \zeta_s (\tilde{\beta}_0 - \beta_0) ds.\end{aligned}$$

Also by definition

$$\partial_w H(w, z) = \partial_w h(w) + \langle \partial_w \rho(w), z \rangle + \frac{1}{2} \langle \partial_w A(w) z, z \rangle + \partial_w r(w, z),$$

and therefore, using that  $r$  is cubic in  $z$  by (H3), we can estimate the factor of  $\epsilon \beta_1$  in the last term in (5.12) by

$$\begin{aligned}\|\partial_w H \circ \zeta_1\|_{\nu-\xi} &\leq \|\partial_w h\|_{\chi-\xi/2} + \|\partial_w \rho\|_{\chi-\xi/2} \|\zeta\|_{\chi-\xi/2} + \frac{1}{2} \|\partial_w A(w)\|_{\chi-\xi/2} \|\zeta\|_{\chi-\xi/2}^2 \\ &\quad + \frac{1}{6} \|\partial_z^3 \partial_w r\|_{\chi-\xi/2, \xi/2} \|\zeta\|_{\chi-\xi/2}^3 \\ &\leq C'_h + \frac{2K\delta^2}{\xi} + \frac{1}{2} \frac{2C_A}{\xi} (K\delta)^2 + \frac{2C_r}{\kappa^3 \xi} (K\delta)^3 \\ &\leq C'_h + \delta + \frac{1}{2} C_A K \delta + \frac{C_r (K\delta)^2}{\kappa^3}.\end{aligned}\quad (5.13)$$

Here we have used Lemma 4.1, Corollary 5.1, the first condition in (5.6a) and Cauchy estimates. We then estimate  $\rho_+$ . From (5.12) and (5.5b) we obtain

$$\begin{aligned}\|\rho_+\|_{\nu-\xi} &\leq \epsilon \|\partial_z^2 H\|_{\nu-\xi/2, \xi/2} \|\partial_w \zeta\|_{\nu-\xi/2} \|\beta_0\|_{\nu-\xi/2} (1 + \epsilon \|\partial_w \zeta\|_{\nu-\xi/2} \|\beta_1\|_{\nu-\xi/2}) \\ &\quad + \epsilon \|\partial_w H \circ \zeta_1\|_{\nu-\xi} \|\beta_1\|_{\nu-\xi/2} \\ &\leq \epsilon 2\kappa^{-2} C_H \frac{2K\delta}{\xi} \frac{4C_D K \delta}{\xi} \left(1 + \epsilon \kappa^{-1} \frac{2K\delta}{\xi} \frac{4C_D K \delta}{\xi}\right) + \epsilon \kappa^{-1} \|\partial_w H \circ \zeta_1\|_{\nu-\xi} \frac{4C_D K \delta}{\xi} \\ &\leq 4\epsilon C_D K (2\kappa^{-2} C_H (1 + 2\epsilon C_D \kappa^{-1}) + \kappa^{-1} \|\partial_w H \circ \zeta_1\|_{\nu-\xi}) \xi^{-1} \delta\end{aligned}$$

Now, introduce

$$c_\delta = 4C_D K \left( 2\kappa^{-2} C_H (1 + 2\epsilon C_D \kappa^{-1}) + \kappa^{-1} \left( C'_h + \delta + \frac{1}{2} C_A K \delta + \frac{C_r (K\delta)^2}{\kappa^3} \right) \right)$$

so that (5.6e) holds true.

For  $A_+$  we have using (5.6b):

$$\begin{aligned}
A_+(w_+) &= \partial_{z_+}^2 H_+(w_+, 0) = \partial_z^2 H(\partial_{z_+} z)^2 + \partial_w^2 H(\partial_{z_+} w)^2 + 2\partial_z \partial_w H(\partial_{z_+} z)(\partial_{z_+} w) \\
&\quad + \partial_z H \partial_{z_+}^2 z + \partial_w H \partial_{z_+}^2 w \\
&= \partial_z^2 H(I + \epsilon \partial_w \zeta_1 \tilde{\beta}_1)^2 + \epsilon^2 \partial_w^2 H \beta_1^2 + 2\epsilon \partial_z \partial_w H(I + \epsilon \partial_w \zeta_1 \tilde{\beta}_1) \beta_1 \\
&\quad + \epsilon \partial_z H(\partial_w^2 \zeta_1 (\tilde{\beta}_1)^2 + \partial_w \zeta_1 \tilde{\beta}_2(w_+, 0) + \epsilon \partial_w H \beta_2(w_+, 0)).
\end{aligned}$$

Here we used (5.3a), (5.3b) and (5.3c), and all derivatives of  $H$  on the right hand sides are evaluated at  $(w, z)(w_+, 0) = (\underline{w}_+, \zeta_1)$ , cf (5.3a), (5.7), (5.9). Using  $\partial_z^2 H(\underline{w}_+, \zeta_1) = A(\underline{w}_+) + \partial_z^2 r(\underline{w}_+, \zeta_1)$  gives, with (5.7),

$$\begin{aligned}
A_+(w_+) - A(w_+) &= \epsilon \int_0^1 \partial_w A(w_+ + \epsilon s \beta_0) \beta_0 ds + \partial_z^2 r \\
&\quad + 2\epsilon \partial_z^2 H(\partial_w \zeta_1 \tilde{\beta}_1) + \epsilon^2 \partial_z^2 H(\partial_w \zeta_1 \tilde{\beta}_1)^2 \\
&\quad + \epsilon^2 \partial_w^2 H \beta_1^2 + 2\epsilon \partial_z \partial_w H(I + \epsilon \partial_w \zeta_1 \tilde{\beta}_1) \beta_1 \\
&\quad + \epsilon \partial_z H(\partial_w^2 \zeta_1 (\tilde{\beta}_1)^2 + \partial_w \zeta_1 \tilde{\beta}_2) + \epsilon \partial_w H \beta_2.
\end{aligned} \tag{5.14}$$

Since  $r$  is cubic in  $z$  by (H3) we have  $\partial_z^2 r(w, z) = \frac{1}{2} \int_0^1 (1-s)^2 \partial_z^3 r(w, sz) z ds$ . Therefore, applying Cauchy-estimates on  $\partial_w A$  and  $\partial_z^3 r$  and using Lemma 5.1, (5.5a), (5.5b), and Lemma 4.1 we obtain

$$\begin{aligned}
\|A_+ - A\|_{\nu-\xi} &\leq \epsilon \frac{2C_A}{\xi} \frac{4C_D K \delta}{\xi} + \frac{6C_r}{\kappa^3} K \delta \\
&\quad + \epsilon \frac{2C_H}{\kappa^2} \left( 2 \frac{2K\delta}{\xi} \frac{4C_D K \delta}{\kappa \xi} + \epsilon \left( \frac{2K\delta}{\xi} \frac{4C_D K \delta}{\kappa \xi} \right)^2 \right) \\
&\quad + \epsilon^2 \frac{2\|\partial_w H \circ \zeta_1\|_{\nu-\xi}}{\xi} \left( \frac{4C_D K \delta}{\kappa \xi} \right)^2 + 2\epsilon \frac{2}{\kappa} \|\partial_w H \circ \zeta_1\|_{\nu-\xi} \left( 1 + \epsilon \frac{2K\delta}{\xi} \frac{4C_D K \delta}{\kappa \xi} \right) \frac{4C_D K \delta}{\kappa \xi} \\
&\quad + \epsilon \frac{C_H}{\kappa} \left( \frac{8K\delta}{\xi^2} \left( \frac{4C_D K \delta}{\kappa \xi} \right)^2 + \frac{2K\delta}{\xi} \frac{8C_D K \delta}{\xi} \right) + \epsilon \|\partial_w H \circ \zeta_1\|_{\nu-\xi} \frac{8C_D K \delta}{\kappa^2 \xi},
\end{aligned} \tag{5.15}$$

so that, by (5.13) and conditions (5.6a),

$$\|A_+ - A\|_{\nu-\xi} \leq c_{A,1} \delta + \frac{\epsilon c_{A,2} \delta}{\xi^2} \tag{5.16}$$

for  $\epsilon$  sufficiently small. Here

$$c_{A,1} = \frac{6C_r K}{\kappa^3},$$

and

$$\begin{aligned}
c_{A,2} &= 8C_A C_D K + \frac{32C_H C_D K^2 \delta}{\kappa^3} + \frac{32\epsilon C_h C_D^2 K^2}{\kappa^4} \delta + \frac{32C_H C_D^2 K}{\kappa^3} + \frac{16C_H C_D K^2 \delta}{\kappa} \\
&\quad + \left( \frac{16\epsilon C_D^2 K}{\kappa^2} + \frac{8C_D K \nu}{\kappa^2} \left( 1 + \frac{2C_D}{\kappa} \right) + \frac{16C_D K \nu}{\kappa^2} \right) \|\partial_w H \circ \zeta_1\|_{\nu-\xi},
\end{aligned}$$

and we use that  $\xi \leq \nu$  and  $\xi \geq 2K\delta$ . The rows in (5.15) correspond to the rows of  $A_+ - A$  as they appear in (5.14). As a consequence we obtain

$$\frac{2}{K}\|y\| \leq \|A_+\|\|y\| + \left(\frac{\epsilon c_{A,1}}{\xi^2} + c_{A,2}\right)\delta\|y\|,$$

and so, if we choose  $\epsilon > 0$  small enough such that

$$\frac{K}{2} \left( \frac{\epsilon c_{A,1}}{\xi^2} + c_{A,2} \right) \delta < 1, \quad (5.17)$$

bearing in mind that  $\delta = O(\epsilon)$  by (H2), then

$$\frac{2}{K} \left( 1 - \frac{K}{2} \left( \frac{\epsilon c_{A,1}}{\xi^2} + c_{A,2} \right) \delta \right) \|y\| \leq \|A_+\|\|y\|,$$

so that (5.6d) holds true. Now finally for  $r_+$ , using (5.6b), (5.1) and (5.3a) we obtain

$$\begin{aligned} r_+(w_+, z_+) &= H(w, z) - \left( h_+(w_+) + \langle \rho_+(w_+), z_+ \rangle + \frac{1}{2} \langle A_+(w_+) z_+, z_+ \rangle \right) \\ &= h(w) - h_+(w_+) \\ &\quad + \langle \rho(w) - \rho_+(w_+), z_+ \rangle + \langle \rho(w), \zeta(\tilde{w}) \rangle \\ &\quad + \frac{1}{2} \langle (A(w) - A_+(w_+)) z_+, z_+ \rangle \\ &\quad + \langle A(w) \zeta(\tilde{w}), z_+ + \frac{1}{2} \zeta(\tilde{w}) \rangle + r(w, z), \end{aligned} \quad (5.18)$$

Here we used that

$$\frac{1}{2} \langle A z_+, z_+ \rangle - \frac{1}{2} \langle A z, z \rangle = \langle A(z_+ - z), \frac{1}{2}(z_+ + z) \rangle$$

We now write  $h(w) - h_+(w_+)$  as  $(h(w) - h(w_+)) + (h(w_+) - h_+(w_+))$ , and, using (5.3a), write

$$h(w) - h(w_+) = h(w_+ + \epsilon\beta) - h(w_+) = \int_0^1 \partial_w h(w_+ + s\epsilon\beta) \epsilon\beta ds.$$

Similarly, we split  $A(w) - A_+(w_+)$  and  $r(w, z) - r_+(w_+, z_+)$ , writing

$$A(w) - A(w_+) = A(w_+ + \epsilon\beta) - A(w_+) = \epsilon \int_0^1 \partial_w A(w_+ + s\epsilon\beta) \beta ds,$$

and

$$r(w, z) - r(w_+, z_+) = r(w, z) - r(w, z_+) + r(w, z_+) - r(w_+, z_+).$$

We then use (5.6c) to estimate  $A(w_+) - A_+(w_+)$ . Using (5.11), (5.5a), Lemma 5.1, Lemma 4.1 and (5.6e) we obtain

$$\begin{aligned}
\|r_+ - r\|_{\nu-\xi, \sigma-\xi} &\leq c_{h,1}\delta + c_{h,2}\frac{\epsilon\delta}{\xi^2} + \epsilon C'_h \frac{4C_D K \delta}{\xi} \\
&\quad + C_D(\delta + \delta_+) + K\delta^2 \\
&\quad + \frac{1}{2}\epsilon \frac{2C_A}{\xi} \frac{4C_D K \delta}{\xi} C_D^2 + \frac{1}{2} \left( \frac{\epsilon c_{A,1}\delta}{\xi^2} + c_{A,2}\delta \right) C_D^2 \\
&\quad + C_A K \delta (C_D + \frac{1}{2}K\delta) + \frac{C_r K \delta}{\kappa} + \epsilon \frac{8C_r C_D K \delta}{\xi^2} \\
&\leq c_{r,1}\delta + \frac{c_{r,2}\epsilon\delta}{\xi^2},
\end{aligned} \tag{5.19}$$

where

$$\begin{aligned}
c_{r,1} &= c_{h,1} + C_D + K\delta + \frac{1}{2}c_{A,2}C_D^2 + C_A C_D K + \frac{1}{2}C_A K^2 \delta + \frac{C_r K}{\kappa}, \\
c_{r,2} &= c_{h,2} + 4C_D K C'_h \nu + C_D c_\delta \nu + 4C_D^3 C_A K + \frac{1}{2}c_{A,1}C_D^2 + 8C_r C_D K.
\end{aligned}$$

As above, the rows in the estimate of  $r_+ - r$  in (5.19) correspond to the rows of  $r_+$  as they appear in (5.18).

Finally, we note that (5.11) gives

$$\|h_+\|_{\nu-\xi} \leq C_{h_+} = C_h + c_{h,1}\delta + c_{h,2}\frac{\epsilon\delta}{\xi}$$

and

$$\begin{aligned}
\|\partial_w h_+\|_{\nu-3\xi/2} &\leq \|\partial_w(h_+ - h)\|_{\nu-3\xi/2} + \|\partial_w h\|_{\nu-\xi/2} \\
&\leq C'_{h_+} = C'_h + \frac{2c_{h,1}\delta}{\xi} + \frac{2c_{h,2}\epsilon\delta}{\xi^2},
\end{aligned}$$

using a Cauchy estimate. ■

Next, we assume that (H4) holds true. The following lemma shows that the equilibrium is contained within the improved slow manifold  $\{z_+ = 0\}$ .

**Lemma 5.4** *Assume (H4). In the  $(w_+, z_+)$ -coordinates the equilibrium is given as  $(w_+^e, z_+^e) = (w^e, 0)$ . □*

**PROOF** First note that  $z^e = 0$  implies that  $\zeta(w^e) = 0$ . By the definition of the symplectic transformation  $\Psi_+$  from the Iterative Lemma 5.3 and the local uniqueness of the equilibrium, assumed in (H4), it will suffice to show that  $(w_+^e, z_+^e) = (w^e, 0)$  satisfies  $\Psi_+(w_+^e, z_+^e) = (w_+^e, z_+^e)$ , i.e. solves the following system:

$$\begin{aligned}
z_+^e &= z^e - \zeta(u^e, v_+^e) = -\zeta(u^e, v_+^e), \\
u_+^e &= u^e + \epsilon \langle \partial_{v_+} \zeta_y, x^e \rangle - \epsilon \langle \partial_{v_+} \zeta_x, y_+^e \rangle = u^e - \epsilon \langle \partial_{v_+} \zeta_x, y_+^e \rangle, \\
v_+^e &= v^e - \epsilon \langle \partial_u \zeta_y, x^e \rangle + \epsilon \langle \partial_u \zeta_x, y_+^e \rangle = v^e + \epsilon \langle \partial_u \zeta_x, y_+^e \rangle.
\end{aligned}$$

Insertion shows this directly. ■

To prove Theorem 3.2 we proceed as in the proof of Theorem 3.1: We will first apply the transformation:

$$\begin{aligned} \Psi_+ : (\mathcal{V} + i(\nu_0 - \xi_0)) \times (\mathcal{S} + i(\sigma_0 - \xi_0)) &\rightarrow (\mathcal{V} + i(\nu_0 - \xi_0/2)) \times (\mathcal{S} + i(\sigma_0 - \xi_0/2)), \\ (w_1, z_1) &\mapsto (w_0, z_0), \end{aligned}$$

once with  $\xi = \xi_0 = \mathcal{O}(1) < \max(\nu_0 - \underline{\nu}, \sigma_0 - \underline{\sigma})$  large. By the Iterative Lemma 5.3, using that  $\|\delta_0\| = \mathcal{O}(\epsilon)$  by (H2), we can therefore take  $\delta_1 = C_\delta \epsilon^2$ , as in (4.16), and start from  $(\mathcal{V} + i\nu_1) \times (\mathcal{S} + i\sigma_1)$  with  $\nu_1 = \nu_0 - \xi_0$  and  $\sigma_1 = \sigma_0 - \xi_0$ . Also,  $\|h_1 - h_0\|_{\nu_1}$ ,  $\|A_1 - A_0\|_{\nu_1}$ ,  $\|r_1 - r_0\|_{\nu_1, \sigma_1} = \mathcal{O}(\epsilon)$ . We then continue to apply the transformations

$$\begin{aligned} \Psi_+ : (\mathcal{V} + i\nu_{n+1}) \times (\mathcal{S} + i\sigma_{n+1}) &\rightarrow (\mathcal{V} + i(\nu_n - \xi_n/2)) \times (\mathcal{S} + i(\sigma_n - \xi_n/2)), \\ (w_{n+1}, z_{n+1}) &\mapsto (w_n, z_n), \end{aligned}$$

successively with  $\xi_n = 2\epsilon \max\{c_{\delta_n}, K_n\}$  for  $n \geq 2$  and  $\xi_1 = 2\epsilon \max\{c_{\delta_1}, K_1, \sqrt{2K_1 C_D C_\delta}\}$ , so that Lemma 5.2 applies and

$$\delta_{n+1} \leq \frac{1}{2} \delta_n, \quad (5.20)$$

for  $1 \leq n \leq N$ . Here  $\delta_n = \|\rho_n\|_{\nu_n}$  with  $\nu_n = \nu_1 - \sum_{i=1}^{n-1} \xi_i$ ,  $\sigma_n = \sigma_1 - \sum_{i=1}^{n-1} \xi_i$ . Then  $\delta_{n+1} \leq 2^{-n} \delta_1$ . To control the possible growth of the constants  $C_{h_n}, C'_{h_n}, C_{A_n}, C_{r_n}, K_n$  with  $n$  we may first restrict  $N$  not only by the conditions that the domains remain non-empty:  $\nu_N - \xi_N \geq \underline{\nu}$  and  $\sigma_N - \xi_N \geq \underline{\sigma}$ , but also by

$$C_{h_n}, C'_{h_n}, C_{A_n}, C_{r_n}, K_n \leq 2 \max\{C_{h_1}, C'_{h_1}, C_{A_1}, C_{r_1}, K_1\} \quad \text{for all } n \leq N. \quad (5.21)$$

Then we can estimate the constants  $c_{h,1}, c_{h,2}, c_\delta, c_{A,1}, c_{A,2}$  and  $c_{r,1}, c_{r,2}$  in the Iterative Lemma 5.3 uniformly with respect to  $n \leq N$ . This allow us to estimate the growth of the constants  $C_{h_n}, C'_{h_n}, C_{A_n}, C_{r_n}$ , (summing over  $n$ , using  $\delta_n = 2^{-n+1} \delta_1$ ,  $\delta_1 = \mathcal{O}(\epsilon^2)$  and the geometric series formula, similarly as for the estimation of  $\beta_n$  in the proof of Theorem 3.1, see (4.26), (4.27)). Moreover we estimate  $K_n$  using (5.6d), similarly as we did for  $\alpha_n$  in (4.22), (4.23). From this we are directly led to the conclusion that the conditions (5.21) pose no restrictions on  $N$ , as they can be satisfied uniformly in  $n$  by choosing  $\epsilon$  sufficiently small. Furthermore,  $\xi_n \leq \bar{\xi}$  for some  $\bar{\xi} = \mathcal{O}(\epsilon)$  giving  $N = \mathcal{O}(\epsilon^{-1})$  possible step. Since  $\delta_1 = \mathcal{O}(\epsilon^2)$  we finally conclude that there exist new constants so that

$$\begin{aligned} \|h_{N+1} - h_1\|_{\underline{\nu}} &\leq \bar{c}_h \epsilon^2, \quad \|\rho_{N+1}\|_{\underline{\nu}} \leq \bar{c}_\delta \epsilon^2 e^{-c/\epsilon}, \\ \|A_{N+1} - A_1\|_{\underline{\nu}} &\leq \bar{c}_A \epsilon, \quad \|r_{N+1} - r_1\|_{\underline{\nu}, \underline{\sigma}} \leq \bar{c}_r \epsilon, \end{aligned}$$

for  $\epsilon \leq \bar{\epsilon}$ . This completes the proof of Theorem 3.2.

## 6 An invariant two-dimensional slow manifold

A consequence of Theorem 3.2 is the persistence of a two dimensional normally elliptic slow manifold with exponentially small gaps. We will sketch this result here which extends a

result in [9] to several fast variables. Consider a real analytic slow-fast Hamiltonian system  $H = H(w_0, z_0)$  with 1 slow degree of freedom,  $d_{\mathcal{Z}}$  fast degrees of freedom and assume that  $z_0 = 0$  is invariant for  $\epsilon = 0$  for  $w \in \mathcal{V} \subset \mathbb{R}^2$ . We assume that  $\{z_0 = 0, \epsilon = 0\}$  is filled with a non-degenerate family of periodic orbits. Then by Theorem 3.2 there exists a transformation  $(w, z) \mapsto (w_0, z_0)$  that transforms  $H$  system into

$$H = h(w) + \langle \rho(w), z \rangle + \frac{1}{2} \langle A(w)z, z \rangle + r(w, z),$$

with  $\rho = \mathcal{O}(e^{-c/\epsilon})$  provided  $\epsilon$  is sufficiently small. The Hamiltonian system  $h = h(w)$  is integrable since it is a one-degree of freedom system. By assumption we can therefore introduce action angle variables  $(\phi, I) \mapsto w$  so that  $h = h(I)$  with  $\partial_I h \neq 0$ . This transformation does not depend upon the fast variables and can therefore directly be lifted to the full space to give:

$$H = h(I) + \langle \tilde{\rho}(\phi, I), z \rangle + \frac{1}{2} \langle \tilde{A}(\phi, I)z, z \rangle + \tilde{r}(\phi, I, z),$$

where  $\tilde{\rho}(\phi, I) = \rho(w)$ ,  $\tilde{A}(\phi, I) = A(w)$  and  $\tilde{r}(\phi, I, z) = r(w, z)$ . We will henceforth suppress the dependency on  $I$ . Since  $\partial_I h \neq 0$  we can solve the equation  $H = E$  of energy conservation for  $I = K_E(\phi, z)$  when  $z$  is small. Differentiating  $H(\phi, K_E, z) = E$  with respect to  $z$  gives  $\partial_z K_E = -(\partial_I H)^{-1} \partial_z H$  and therefore:

$$\epsilon \partial_\phi z = -J_z \partial_z K_E,$$

where  $J_z$  is the matrix from (3.5). Here

$$K_E = h^{-1}(E) - \langle \rho_E, z \rangle - \frac{1}{2} \langle A_E z, z \rangle - r_E,$$

with  $\rho_E = (\partial_I h)^{-1} \tilde{\rho} = \mathcal{O}(e^{-c/\epsilon})$ ,  $A_E = (\partial_I h)^{-1} (\tilde{A} + \mathcal{O}(e^{-c/\epsilon}))$  and  $r_E = \mathcal{O}(z^3)$ . Fix  $\epsilon$  small and introduce  $\mu^2 = \sup \rho_E = \mathcal{O}(e^{-c/\epsilon})$  so that  $\rho_E = \mu^2 \hat{\rho}$  with  $\sup \hat{\rho} = 1$ . Also let  $\mu^2 B = A_E - \tilde{A}$ ,  $\psi = (\partial_I h)^{-1} \phi$  and  $z = \mu \bar{z}$ . We then introduce

$$Q = Q(\psi, \bar{z}) = \mu \langle \hat{\rho}, \bar{z} \rangle + \langle (\tilde{A} + \mu^2 B) \bar{z}, \bar{z} \rangle + \mu \bar{r},$$

$\bar{r}(\phi, \mu \bar{z}) = \mu^3 r_E(\psi, \bar{z})$  so that

$$\epsilon \partial_\psi \bar{z} = J_z \partial_z Q = \mu J_z (\hat{\rho} + \mu B \bar{z} + \partial_{\bar{z}} \bar{r}) + J_z A \bar{z}.$$

Let

$$P_\mu : \{\psi = 0\} \rightarrow \{\psi = T\}, T = 2\pi(\partial_I h)^{-1},$$

be the corresponding stroboscopic mapping. It is symplectic since the system is Hamiltonian. Also  $P_0(0) = 0$ . The persistence of this fixed point for  $\mu \neq 0$  provides the persistence of the periodic orbits, which we have parametrized by  $E$ . For this we consider the linear map  $\partial_{\bar{z}} P_0 = \Psi(T)$ , the monodromy matrix, where  $\Psi = \Psi(\psi)$  satisfying  $\Psi(0) = I$ , is the fundamental matrix of the system

$$\epsilon \partial_\psi \bar{z} = J_z A \bar{z}.$$

The eigenvalues of  $\Psi(T)$ ,  $\lambda_1(E), \dots, \lambda_{2d_Z}(E)$ , are the characteristic multipliers and they depend upon  $E$ . We can apply the implicit function theorem to the set

$$\mathcal{E} = \{E \mid |\lambda_i(E) - 1| \geq \sqrt{\mu}, i = 1, \dots, 2d_Z\}, \quad (6.1)$$

since the perturbation is  $\mathcal{O}(\mu) = \mathcal{O}(e^{-c/(2\epsilon)})$ . Generically one would expect the singular set  $\{E \mid |\lambda_i(E) - 1| = 0 \text{ for some } i\}$  to be discrete containing simple solutions and without accumulation points. In the affirmative case the complement of the set  $\mathcal{E}$  from (6.1) would be exponentially small, and the slow manifold  $\{z_0 = 0, \epsilon = 0\}$  would therefore perturb to an invariant slow manifold  $\{z_0 = \mathcal{O}(\epsilon)\}$ ,  $\epsilon \leq \epsilon_0$  with exponentially small gaps. Note that this result does not depend upon resonances between the fast variables.

## 7 Conclusion

We proved the existence of slow manifolds which are invariant up to an exponentially small error in general and Hamiltonian analytic slow-fast systems. The slow manifolds were constructed so that they included nearby equilibria. The approach we used for general systems is due to R. S. MacKay and does not require the transformation of the slow variables, and as a consequence, our results in this case hold true for unbounded fast vector-fields. In future work we aim to extend our results to unbounded slow vector-fields and infinite dimensional slow-fast Hamiltonian systems. Here we believe that there is great potential in combining the methods used here with the methods due to K. Matthies and A. Scheel [19].

## Acknowledgement

The authors would like to thank R. M. Roberts for his help with preparing the manuscript. K. Uldall Kristiansen was funded by the H. C. Ørsted post doc stipend grant.

## References

- [1] R. Bellmann. The stability of solutions of linear differential equations. 10:643–647, 1981.
- [2] J. Carr. *Applications of centre manifold theory*, volume 35. New York: Springer-Verlag, 1981.
- [3] C. J. Cotter and S. Reich. Adiabatic invariance and applications to MD and NWP. *BIT Numerical Mathematics*, 3:439–455, 2003.
- [4] N. Fenichel. Persistence and smoothness of invariant manifolds for flows. *Indiana University Mathematics Journal*, 21:193–226, 1971.
- [5] N. Fenichel. Asymptotic stability with rate conditions. *Indiana University Mathematics Journal*, 23:1109–1137, 1974.
- [6] S.J. Fraser. The steady state and equilibrium approximations: a geometrical picture. *Journal of Chemical Physics*, 88:4732–4738, 1990.



- [7] C. W. Gear, T. J. Kaper, I. G. Kevrekidis, and A.M. Zagaris. Projecting to a slow manifold: Singularly perturbed systems and legacy codes. *SIAM Journal of Applied Dynamical Systems*, 4(3):711–732, 2005.
- [8] V. Gelfreich and L. Lerman. Almost invariant elliptic manifold in a singularly perturbed Hamiltonian system. *Nonlinearity*, 15:447–557, 2002.
- [9] V. Gelfreich and L. Lerman. Long-periodic orbits and invariant tori in a singularly perturbed Hamiltonian system. *Physica D*, 176:pp 125–146, 2003.
- [10] Dan Henry. *Geometric Theory of Semilinear Parabolic Equations*. Lecture Notes in Mathematics. Springer Verlag, Berlin, Heidelberg, New York, 1981.
- [11] J. Pöschel, and E. Trubowitz . *Inverse Spectral Theory*, volume 130. Pure and Applied Mathematics, Academic Press, Inc., 1987.
- [12] H. G. Kaper and T. J. Kaper. Asymptotic analysis of two reduction methods for systems of chemical reactions. *Physica D*, 165:66–93, 2002.
- [13] J. Laskar. Large scale chaos in the solar system. *Astronomy and Astrophysics*, 287, 1994.
- [14] J. Laskar and M. Gastineau. Existence of collisional trajectories of Mercury, Mars and Venus with the Earth. *Nature*, 459, 11 June 2009.
- [15] E. N. Lorenz. The slow manifold - what is it? *American Meteorological Society*, 15 December, 1992.
- [16] E. N. Lorenz. Existence of a slow manifold. *Journal of the Atmospheric Sciences*, 43(15):1547–1557, 1986.
- [17] R. S. MacKay. Slow manifolds. In: “*Energy Localisation and Transfer*”, eds T Dauxois, A Litvak-Hinzenzon, RS MacKay, A Spanoudaki, *World Scientific*, pages 149–192, 2004.
- [18] R. Mañé. Persistent manifolds are normally hyperbolic. *Bulletin of American Mathematical Society*, 80:90–91, 1980.
- [19] K. Matthies and A. Scheel. Exponential averaging for Hamiltonian evolution equations. *Transactions of the American Mathematical Society*, 355(2):747–773, 2002.
- [20] D. A. McQuarrie. *Physical Chemistry: A Molecular Approach*. Sausalito: University Science Books, 1997.
- [21] L. Michaelis and M. Menten. Die Kinetik der Invertinwirkung. *Biochemische Zeitschrift*, 49:333–369, 1913.
- [22] J. D. Murray. *Mathematical Biology*, volume 19. Berlin: Springer-Verlag, 1993.
- [23] A. Neishtadt. On the accuracy of conservation of adiabatic invariant. *Journal of Applied Mathematics and Mechanics*, 45(1):58–63, 1982.

- [24] A. Neishtadt. The separation of motions in systems with rapidly rotating phase. *Journal of Applied Mathematics and Mechanics*, 48(2):133–139, 1984.
- [25] V. Rom-Kedar and D. Turaev. The symmetric parabolic resonance instability. *Nonlinearity*, 23:1325–1351, 2010.
- [26] R. Temam. Inertial Manifolds. *The Mathematical Intelligencer*, 12(4):68–74, 1990.
- [27] K. Uldall Kristiansen, P. Palmer, and R. M. Roberts. A unification of models of tethered satellites. *SIAM Journal of Applied Dynamical Systems*, 10:1042–1069, 2011.
- [28] K. Uldall Kristiansen, P. Palmer, and R. M. Roberts. The persistence of a slow manifold with bifurcation. *SIAM Journal of Applied Dynamical Systems*, 11:661–683, 2012.
- [29] J. Vanneste. Asymptotics of a slow manifold. *SIAM Journal of Applied Dynamical Systems*, 7, 2008.
- [30] D. Wirosoetisno. Exponentially accurate balance dynamics. *Advances in Differential Equations*, 9, 2004.
- [31] A.M. Zagaris, C. W. Gear, T. J. Kaper, and I. G. Kevrekidis. Analysis of the accuracy and convergence of equation-free projection to a slow manifold. *ESAIM: Mathematical Modelling and Numerical Analysis*, 43:757–784, 2009.